

Motion Analysis Marker for Ease of Image Processing

BHARTENDU SETH

Department of Mechanical Engineering
Indian Institute of Technology Bombay

Mumbai 400 076,

INDIA

sethbharat@gmail.com <http://www.me.iitb.ac.in/instruction/faculty/seth.html>

Abstract: High speed machines and mechanisms are often studied from a sequence of images obtained from high speed videography. The use of markers printed or attached on moving parts can greatly assist in tracking the moving parts. We present the design of a marker suitable for planar motion analysis of mechanism. The marker is designed to make the task of image processing computationally less intensive so as to make real time motion analysis practical. Rosenfeld equivalence table algorithm is used to segment the input image. The new geometry of marker facilitates automatic tracking and provides both position and orientation information. Hu invariant moments are used for recognition of the marker shape in the segmented image. Markers are uniquely identified on the basis of a text code that is placed in a designated location with respect to the marker geometry. In this method, the bounding box for the text code is computed. Knowing the orientation of the marker and therefore the text orientation, it is possible to transform the sub-image, containing the text, so that the text is aligned horizontally. This will permit a standard OCR routine to read the text code. The motion of various moving parts in image sequence can be easily inferred once the position and orientation of each of the marker is known.

Key words: image processing, asymmetric marker, video-graphic motion tracking, planar mechanism, occlusion, position and orientation determination

1. Introduction

Image processing techniques are being used in a wide spectrum of problems, such as, optical character recognition (OCR) of scanned documents, DNA sequence analysis in bioinformatics, inspection and motion analysis related to industrial automation, personnel recognition from images of face, fingerprints or eyes, crop estimations using remote satellite images and speech recognition. While in many cases the regions of interest have to be recognized in the image from the colour, texture or other pattern information, in some situations, such as machine parts or mechanisms, it is permissible to use special identification tags, known as markers to facilitate the identification of the points of interest. In applications requiring precise coordinated motion at high speeds, video imaging techniques are indispensable for ensuring that the motion requirements are actually satisfied under the actual operating conditions.

Most commercially available motion tracking software, such as Autotrack™ [1], use checkerboard markers to identify key locations on links of interest. Since checkerboard markers do not give unambiguous orientation information, each moving rigid link, undergoing arbitrary planar motion, requires two or more markers. Moreover, during the operation of a mechanism, some markers are occluded by other moving links, making it necessary to use additional markers for a link. If links undergo deformation, then accurate observation will require even more number of markers. Thus the number of markers in complex mechanisms tends to increase and the problem of quickly and correctly identifying a marker also acquires greater significance. Text coded markers can be an effective way to address the problem of positive identification of multitude of markers.

1.2 The Approach

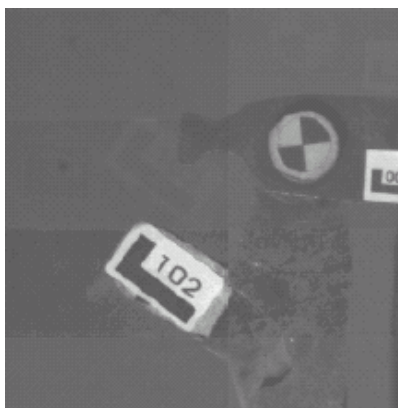


Fig. 1: Conventional and L-shaped markers

The inputs to the system are digitized images of mechanisms such as the one shown in Fig 1, obtained from an appropriate camera and frame grabber system. We need to extract regions in the image that correspond to the markers. In this paper, we have developed a new geometry for markers. The proposed markers are L-shaped with an aspect ratio of two. Our approach here is to use a text identifier that forms an integral part of the marker design. If the two limbs of the marker shape form the axes of a coordinate system, then the text is placed in the first quadrant and generally within the bounding box of the marker. The asymmetry of the shape is used to unambiguously determine the orientation of the. Existing OCR (Optical Character recognition) packages require the text to be horizontally aligned for identification. The orientation of the marker is used for the geometric transformation of the text image region so as to make it compatible for use by available OCR packages like GOCR [2].

2. Background

Automatic determination of motion of links with attached markers from a sequence of images requires several key steps. Starting from image acquisition, the candidate marker regions are distinguished from the other regions of the image through the process of segmentation. The marker segments are then confirmed using shape analysis and their orientations are determined. Finally, the identifying tag of the marker is processed to individually identify each marker. A brief review of work related to these steps is presented below.

2.1 Segmentation

Segmentation is defined as the process of partitioning an entire image into distinct and connected sub-regions, such that the neighboring sub-regions are distinguishable using some property such as colour, grayness or texture. Two approaches used for segmentation are hierarchical representation such as quad-trees and array representation. We use the array representation, in which the image is scanned and neighboring pixels are compared to determine if they belong to the same region or not. A labeling scheme is used to distinguish the different regions. Array approach requires several passes on the image data before the image is segmented. Simple connected neighbor search scheme for segmentation is computationally expensive and has a time complexity $O(N^4)$ for an $N \times N$ image. The processing can be simplified by appropriately thresholding the image before labeling [3] so that potential object pixels can be separated from the background pixels. The algorithm can be made significantly more efficient by the Equivalence Table based approach, proposed by Rosenfeld et al [4], which requires only two passes on the image data and has time complexity of $O(N^2/P)$ for an $N \times N$ image with P distinct segments.

2.2 Recognition

In the template matching recognition approach, similarity between a region and a template are compared after suitably scaling and rotating an image template. This approach is computationally demanding. The main advantage of this approach is that it could also be used without the use of markers. Researchers have also extended the applicability of this approach to rubber sheet deformable templates [5], which makes it applicable to markers that rotate in planes skewed with respect to the camera axis. The statistical approach is based on a measured feature-vector. The decision boundaries are determined by probability distributions function of the patterns belonging to each class, which must either be specified or learned [6]. In the syntactic approach, each pattern is viewed as being composed of simpler sub-patterns, down to elementary level of primitives [7]. In this approach the patterns are viewed as made of primitives, much the same way as sentences are made from words using well defined rules or grammar. This approach is particularly suitable when the patterns have a definite structure and has been applied to analysis

of ECG waveforms and shape analysis of contours [8]. However, this approach is susceptible to noise problems. In noisy images, the computational effort increases drastically due to the necessity of larger training sets. Bunke [9] has advocated using the approach in combination with statistical approach. Neural net approach has been applied using both feed forward network [10] as well as self organizing networks [11]. This approach is less dependent on the domain-specific knowledge due to the self-learning capabilities of neural networks.

2.3 Shape and Orientation Analysis

Once a region of an image is identified as a potential marker, the shape of the region is analyzed for ascertaining it as the marker shape. Several approaches can broadly be classified as based on geometry, based on moments, based on polygon approximation, Fourier transform based, etc. [12]. The geometry based approach uses properties such as area, perimeter, thinness, convexity. Moment based approaches compute moments of the region, which depend on the shape of the region. It is possible to define moments, such as the Hu moments, which remain invariant for a shape under translation and rotation of the shape [13]. Other approaches use description of the boundary shape as turn angle sequences or shape numbers or approximate the essential shapes with polygonal shapes.

3. Marker Design and Image Processing

3.1 Marker Design

The basic marker geometry consists of a L-shape (see Fig. 2), which is the simplest asymmetric shape. The marker we used has the following features:

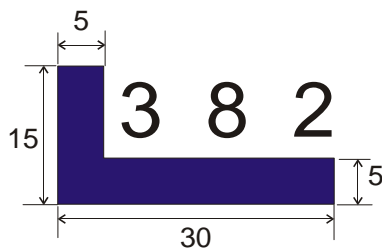


Fig. 2: The first marker prototype (all dimensions in mm)

1. The text-code on the marker has 3 digits – so that the marker range is from 000 to 999.
2. The numbers are printed with a blank character between digits. This aids in OCR segmentation as a low-resolution “13” could be misidentified as “8”.
3. The text font is kept fixed. The OCR system could then be optimized for the font.
4. The text is enclosed within a standard L-shaped marker. The two limbs of the marker act as natural references for the bounding box of the text.
5. The image plane is assumed to be perpendicular to the line of sight of the camera.

3.2 Shape Analysis of Markers using Hu Invariant Moments

We use the Hu invariant moments, based on the fundamental theorem of moment invariants, proposed by Ming-Kei Hu. This theorem states that a particular set of seven moments remain invariant under the operations of rotation and translation. The marker shape is distinctive enough that the first four Hu moments were found to be sufficient to positively identify the markers. The first four normalized Hu moments, are

$$\phi_1 = \eta_{20} + \eta_{02} \tag{1}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{2}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{3}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{4}$$

The normalized central moments, $\eta_{pq} = \mu_{pq} / \mu_{00}^\lambda$, where $\lambda = 1+(p+q)/2$, are invariant with respect to scaling also. μ_{pq} is moment computed relative to the centroid and is given by $\mu_{pq} = \sum_{x,y} (x - \bar{x})^p (y - \bar{y})^q f(x, y)$. Where, $f(x,y)$ is the grayness value at pixel location (x,y) .

3.3 Determining Orientation

Most of the approaches described in literature to determine the orientation use variant central moment method to estimate the rotation of the image segment, which vary only with the rotation of the object. This fact can be used to compute the angle of rotation, θ , given by,

$$\theta = 0.5 \tan^{-1} (2\mu_{11} / (\mu_{20} - \mu_{02}))$$

3.4 Some problems encountered in implementation

Fig. 3 shows the result of segmentation, marker recognition and determination of orientations. It was observed that the variant moments approach failed to give the correct results for the smaller size of the markers as the pixellation problem became significant, rendering the results obtained unreliable. An attempt was made by constructing a

trapezium using the extreme points of the marker to determine the orientation. The base line of the marker was obtained by examining the internal angles of the trapezium. However, reliable results could not be obtained when the marker was almost aligned with the horizontal and vertical. The problem was finally solved by a slight modification to the marker geometry.

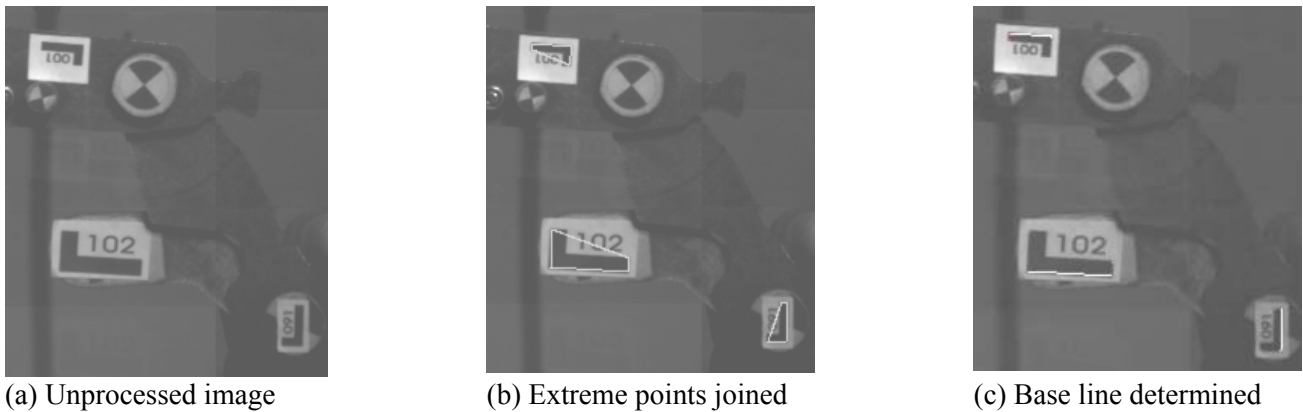


Fig. 3: Marker with bounding-box identified

3.5 Modified Marker Design

The marker was modified as shown in Fig. 4. Orientation of the marker is now determined by finding the corner of the marker at the greatest

Euclidean distance from its centroid location. The modification ensures a greater distance between the farthest corner and the centroid than the other corners.

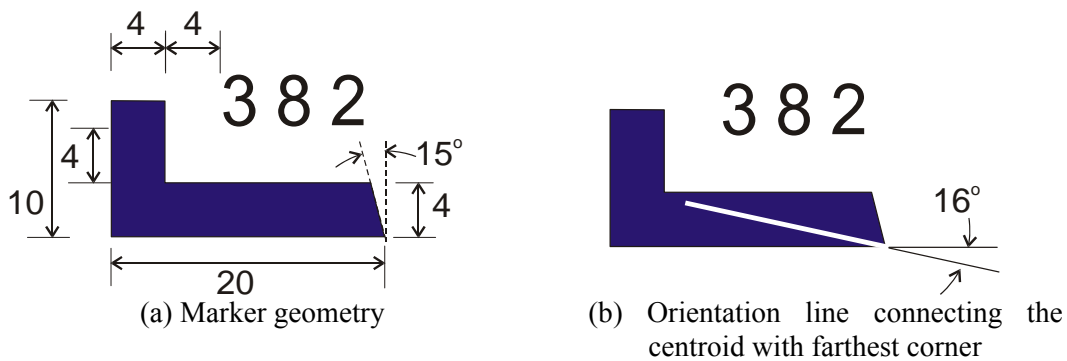


Fig. 4: Modified Marker (all dimensions in mm)

The text box can be determined geometrically as follows:

$$R_{6x} = R_{gx} + l \cos(\theta + 16^\circ) \quad (5)$$

$$R_{6y} = R_{gy} + l \sin(\theta + 16^\circ) \quad (6)$$

$$R_{3x} = R_{6x} - 3/4 l \sin(\theta + 16^\circ) \quad (7)$$

$$R_{3y} = R_{6y} + 3/4 l \cos(\theta + 16^\circ) \quad (8)$$

$$\mathbf{R}_5 = 4/3 \mathbf{R}_g - 1/3 \mathbf{R}_6 \quad (9)$$

$$\mathbf{R}_4 = \mathbf{R}_5 + (\mathbf{R}_3 - \mathbf{R}_6) \quad (10)$$

$$\mathbf{R}_2 = 4/5 \mathbf{R}_6 + 1/5 \mathbf{R}_3 \quad (11)$$

$$\mathbf{R}_1 = 4/5 \mathbf{R}_5 + 1/5 \mathbf{R}_4 \quad (12)$$

where, l is the distance of the farthest corner (point t in Fig. 5) from the centroid (point g). θ is the orientation as defined earlier and points 1,2,3,4 form the bounding box for the text region and the quantities in boldface denote vectors.

Once the bounding box for the text region is obtained along with the angle of rotation, OCR routine can determine the marker number. Fig. 6 illustrates this process of orientation and text box determination for a double pendulum with several markers in different orientation and of two different sizes.

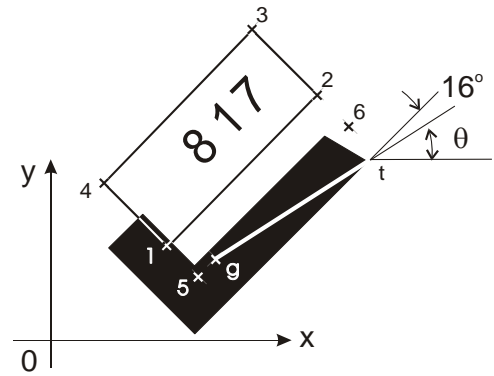


Fig. 5: Location of text bounding box

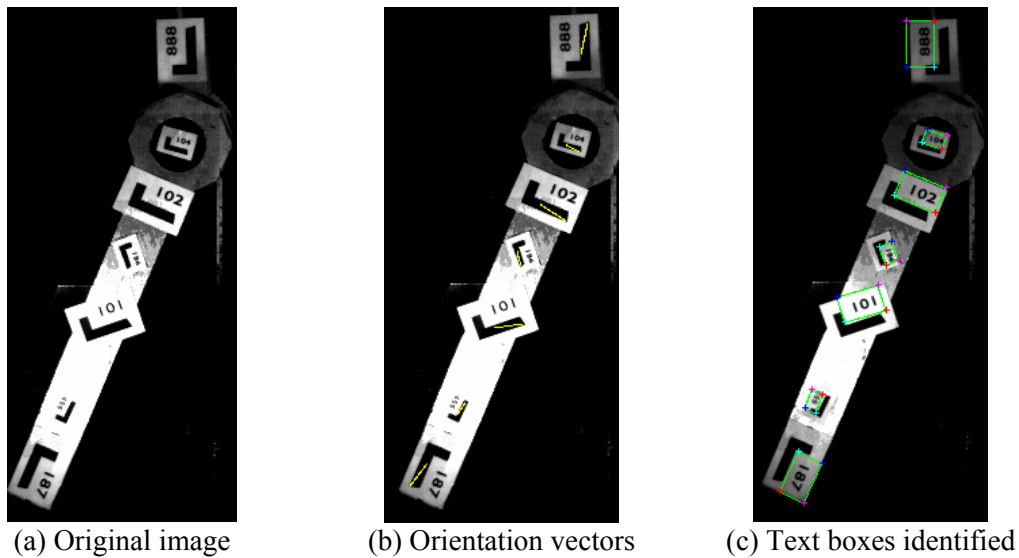


Fig. 6: Determination of marker orientation and text bounding box

4. Results and Conclusions

In the trial runs, the tracking software system using the new marker design performed reliably. Use of Rosenfeld algorithm improved the computation time by two orders of magnitude compared with the 8 connected-neighbor search algorithm. This was estimated using the information available from the Windows Task Manager (processor and memory usage).

For shape analysis, the system was fed with sample images of markers in different orientations and of different sizes. It was found that the first four moments corresponding to the L-shaped marker were in very good agreement with each other, even under scale transformations.

The computation of Invariant moments was speeded up using sub-images based on the maximum expected movement of the markers from one frame to the next. Here, the invariant moments-based algorithm has a time complexity $O((P-1) N^2)$, where P is the number of segments the image is

divided into. In the sub-images both the value of N and P both are smaller and an improvement by a factor of 200 was obtained as compared with the processing time for a 512x512 pixels and 50 segments image.

For orientation-analysis, uneven lighting patterns make it difficult to handle both the bright and dark parts of the image. A Gaussian smoothing step may be added to the system, so as to overcome the problem of uneven lighting and also misidentification of fallacious markers generated by reflection from the camera optics.

The system developed here is well suited for the purpose of motion analysis of mechanisms. The advantage of this system over existing technologies is the ease with which the marker shapes can be identified, thereby speeding up the image processing. Secondly, the human readable text coding uniquely identifies each marker and significantly increases the number of markers that a system using can handle.

References:

- [1] Autotrack v3.01, Vannier-Photoelec, Antony Cedex, France 1998.
- [2] Schulenburg, J., GOCR 0.37, <http://jocr.sourceforge.net/developers.html>, release Aug 11, 2002,
- [3] Rosenfeld A, et al, Fast Pyramidal Algorithms for Thresh-holding, Pattern Recognition, v28, 95.
- [4] Flotz M. A., Connected Components in Binary Images. Machine Vision, 6, 1997.
- [5] Younes L., Optimal Matching between Shapes via Elastic Deformations. Image and Vision Computing, Vol. 17, 5-6, April 1999.
- [6] Yukalov V.I., Probabilistic Approach to Pattern Selection. Physica A: Statistical Mechanics and its Applications, Vol. 291, 1-4, March 2001.
- [7] Tanaka E, Theoretical Aspects of Syntactic Pattern Recognition. Pattern Recognition, Vol. 28, 95.
- [8] Corwell J.R., Syntactic Pattern Recognizer for Vehicle Identification Numbers, Image and Vision Computing, Vol. 13, 1, February 1995.
- [9] Bunke H., Towards bridging the Gap between Statistical and Structural Pattern Recognition: Two New Concepts in Graph Matching. ICAPR – Advances in Pattern Recognition, Second Int'l Conference, March 2001.
- [10] Kraijveld M.A., The Effective Capacity of Multilayer Feedforward Network Classifiers. IEEE, 12th Int'l Conf. on Pattern Recognition, Vol. 2, 1994.
- [11] Kiang M.Y., Extending the Kohonen self-organizing map networks for clustering analysis. Computational Methods and Data Analysis, Vol. 38, 2, December 2001.
- [12] Antani S., Kasturi R., Jain R., A Survey on the use of Pattern Recognition methods for abstraction, indexing and retrieval of images and video. Pattern Recognition 35 (4), 2002.
- [13] Ojala T., A Generalized Local Binary Pattern Operator for Multiresolution Gray Scale and Rotation Invariant Texture Classification. 2nd Int'l Conf. for Advances in Pattern Recognition, March 2001.