Image Based Subpixel Techniques for Movement and Vibration Tracking

David MAS\textsuperscript{1}, Julian ESPINOSA\textsuperscript{1}, Ana Belen ROIG\textsuperscript{2}, Belen FERRER\textsuperscript{2}, Jorge PEREZ\textsuperscript{1}

\textsuperscript{1}I. U. Fisica Aplicada a las Ciencias y Tecnologías, Universidad de Alicante; Alicante, Spain
Ph: +34 965 903400, Fax: +34 965 903464; e-mail: julian.espinoza@ua.es; jorge.perez@ua.es; david.mas@ua.es
\textsuperscript{2}Dep. Optica, Universidad de Alicante, Alicante, Spain; E-mail anabelen.roig@ua.es
\textsuperscript{3}Dep. Ing. Civil, Universidad de Alicante, Alicante, Spain; E-mail belen.ferrer@ua.es

Abstract
Subpixel techniques increase the accuracy and efficiency of image detectors, processing units and algorithms and provide very cost-effective systems for object tracking. Most usual techniques are those based on centroid track or image correlation. We propose here different techniques that can provide some additional advantages depending on the application. From high-speed video sequences, we obtain successful results using object detection through Hough transform, using contour based methods to track ellipses and even without using any target, just with multilevel thresholding.

Keywords: subpixel techniques, vibration tracking, image based measurement

1. Introduction

Measurement of movements and vibrations is a very important task for many sectors, such as assembly of mechanisms, civil structures and buildings monitoring or water level tracking, for example. Among all the procedures available image-based measurements have arisen in the last years due to the increasing calculation capabilities of new computers. Additionally, these systems usually imply non-contact procedures, therefore there is no needed to reach the point to measure and that makes it safer and easier than contact procedures.

In image-based techniques, accuracy of measurements is directly related with the spatial and temporal resolution of the camera that is recording the movement. Unfortunately, cameras with both high temporal and spatial resolution are very expensive and the affordable technology tends to prioritize the acquisition speed at the cost of reducing the image quality.

Therefore, in order to improve the accuracy of measurements on an image, the use of superresolving techniques is needed. In general, tracking methods can only detect movements when the object has moved a complete pixel in the image. Subpixel techniques allow detecting movements smaller than a single pixel. To do so, we propose different methods.

The simplest approach consists of detecting pixel groups that share a common feature and follow this structure along a video sequence. Calculation of the object centroid locates the object in the scene. As far as the common feature remains through all the sequence, the object will be detected and tracked. Sub-pixel resolution arises as a statistical effect since the centroid is calculated taking into account a high number of pixels [1].

The second approach consists of taking the objects that share a common feature and construct an analytical model whose geometrical properties can be accurately calculated [2]. In our case, we introduced elliptical targets so, once detected, its contour can be fitted to the geometrical figure and its centre calculated from the obtained parameters (see figure 1).
The third method we have used in our research consists in detecting differences between two successive frames. A convenient pattern may increase the possibility of this detection even when the movement is below the pixel unit. Theoretically, this can be done using a target with sparse pseudo-random dots. In figure 2, we explain this graphically. Let us now consider that an object with details smaller than one pixel (red dots in figure 2). This detail will not be resolved by the imaging system and its presence will be only revealed by a pixel value different from those from the surrounding (in gray). The location of this detail inside the pixel area is unknown, but if it is close to the pixel border, a small movement in the proper direction will make this detail to change to the next pixel area even with a subpixel shift. A second subpixel shift of the same magnitude will produce a visible effect on the same detail until it gets closer to the next border again. Therefore, another detail will then reveal the subpixel movement [3].
2. Results

2.1 Simple object tracking

The first technique explained is applied to track the movement of a falling object. The experiment consists of an 80 kg cylinder that is dropped on an inclined surface to analyse the capacity of retention of a net located at the end of the inclined surface (fig. 3a). The net is one security device commonly used in building construction to avoid the fall of the workers to the void. Under the impact caused by the cylinder, the net is deformed. The maximum deflection of the net is related to the maximum deceleration of the falling body, so we can calculate the forces on a falling worker and determine the possible injuries on her. As this is a very fast event and the exact trajectory is unknown, tracking the movement from a sequence taken with fast camera was the best solution.

The camera used in this experiment was an AOS X-Pri, working at 500 fps with a frame resolution of 800x560 px. It was located in a lateral view of the ramp and net in order to have also a lateral view of the cylinder. This way the cylinder appears in the image as a circle, which can be identified. To do so, each frame is binarized using an appropriate threshold, and the borders are extracted. In the second step, all dots sharing the contour of a circle of known radius are detected using a Hough transform. In order to maximize the possibilities of detection, the centre of the circle is manually marked in the first frame. After that, the calculated centre in one frame will be used as a seed for the following frame and the process becomes automatic. With the centroid coordinates, the whole trajectory can be reconstructed (see figure 3) and the second derivative will provide the instant acceleration of the ballast. A full explanation of the procedure can be found in [1].

Figure 3. Experiment with security net, (a) general view, (b) trajectory obtained
2.2 Contour fitting

If the objet to be tracked presents a defined shape, its contour can be traced and fitted to an analytical curve. In the case of an ellipse, the dots in the detected contour \((x, y)\) can be fitted to a 6 parameters polynomial:

\[
Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0
\]

From the obtained parameters, the centre position \((x_c, y_c)\), can be determined as:

\[
(x_c, y_c) = \left(\frac{-D}{2A}, \frac{-E}{2C}\right)
\]

Here, the subpixel resolution is achieved by analytical interpolation of the contour line. The method has been shown to provide resolutions below 0.02 px for targets larger than 100 px [2]. This enhanced resolution allows the implementation of this method even with low cost camera. The authors have successfully implemented this method in a wide variety of experiments.

As an example, we show here the measurement of the movement of a pedestrian bridge located inside of a building after an impulse force (figure 4). A person jumping in the middle of the bridge applied an impact load on the bridge. The movement was measured with a Casio Exilim camera that is able to acquire videos at 240 frames per second with a spatial resolution of 512x384 px (figure 5). The camera was located 6 meters away from the structure. In the lateral deck of the bridge a target consisting on a circle of 10 cm (26.5 px) of diameter was attached. Some accelerometers were used as a contrast device for the frequency. The displacement measured was previously calibrated using a linear motorized stage.

Notice that the conversion ratio between object and image sizes is 3.77 mm per pixel. Therefore, the nominal resolution of the system is really low and subpixel resolution is needed in order to provide a reliable measurement above the noise.

![Figure 4. Target in bridge deck](image-url)
Results of the bridge vibration measured by the camera have been depicted in Figure 6. Notice that, despite the noise, the system is able to see vibrations with amplitudes of 0.05 px an below, thus registering oscillations of 0.2 mm. The Fourier transform of the signal from the camera has been compared with that from the accelerometer (see figures 6 and 7), thus showing that our method is capable of registering the main vibrations of the structure.
2.3. Frequency measurement

Finally, we have implemented an application of the third subpixel method described above. As we said, if we have an intricate contour in one shape in the image, we can take advantage of this contour and use it as a “random” target. First task is binarize the image to isolate the object from the background. Any movement in the image leads to a change of illumination, which is revealed in the binarized image. To maximize the probability of detecting the movement, different thresholds can be used and the results given by each one of them can be combined.

Illumination changes can be done by different reasons, not only movement. Additionally, object contour is not exactly a random target with the optimal number of dots. Therefore, for real objects under any kind of movement this procedure is not always successful. However, if the object is vibrating the changes in the image will be periodic and detection of the main frequency will be easier. From the binary images, the analysis consists of counting the variation of active pixels with respect to the first frame and performing a Fourier transform. The full description of this procedure can be found in [4].

This procedure was checked in different experiments that are briefly explained below. In all cases, a CASIO Exilim EX-ZR 1000 pocket camera [5] was used. This camera stores the images in JPEG-AVI format and, although it introduces some noise, it does not affect the results because that noise appears as random dots after binarization. Therefore, this method is able to give results even with a pocket camera. In all experiments, the camera was recording at 1000 fps. For this frame rate, the spatial resolution of the camera is 224x64 px.

The experiment was performed on a loudspeaker membrane. It was forced to sound at only two frequencies (317 and 412 Hz) by a computer. A microphone was also used to record the sound and check our method. In this experiment, a small ROI was taken to make the calculations (figure 8). The results are shown in figure 9. Both camera and microphone find the two imposed frequencies (317 and 412 Hz) and a third one (222 Hz) due to membrane distortion (see [4] for more information).

![Figure 8. ROI taken for the loudspeaker experiment.](image)
Figure 9. Frequencies obtained by microphone and image processing in loudspeaker experiment.

We want to emphasize that this last method is able to measure vibration frequencies without attaching any target or special illumination so is very convenient for many applications where non-invasive techniques are needed. Unfortunately information comes from pixel variation in thresholded levels so there is not a direct connection between the movement amplitude and the peak height.

3. Conclusions

We have presented here three different methods for measuring movements and vibrations with vision systems. Two of the systems are non invasive, while the other only requires the attachment of an adhesive target that can be easily removed. The three systems use subpixel resolution thus enhancing the spatial resolution and allowing their implementation with low cost cameras.

Acknowledgements

The authors acknowledge the support of the Spanish Ministerio de Economía y Competitividad through the project BIA2011-22704, the Generalitat Valenciana through the project PROMETEO/2011/021 and GV/2013/009 and the University of Alicante through the project GRE13-10. A.B. Roig acknowledges the financial support from CajaMurcia

References


