DETECTION AND TRACKING OF SMALL OBJECTS BY DETAILS REMOVAL

S. LEFEVRE, T. FOISSOTTE AND N. VINCENT

Laboratoire d'Informatique, Université de Tours
64 avenue Portalis, 37200 Tours – FRANCE
E-mail: {lefevre,vincent}@univ-tours.fr

In this paper, we deal with very small object tracking, assuming simultaneously moving camera and real time processing constraints. Considering these objects are represented in the images by detail pixels, we propose to segment images into "detail" and "other" pixels, using a multiresolution filtering which allows details removal. A local analysis is then performed to cluster detail pixels into small object candidates. Real small objects are discriminated from noise and tracked through time by integrating the detection obtained on every frame of the video sequence. Some results are finally given to show the efficiency of the proposed method, which is able to track at a high frequency very small objects in a video sequence acquired with a moving camera.

1. Introduction

Due to the growth of multimedia data available, the need for specific multimedia information systems is more and more important and efficients tools are required. Among these tools, object tracking is often necessary to help understand the content of a given scene. As different kinds of objects can be tracked (rigid or non-rigid objects, small or large objects, *etc*), specific tracking algorithms have been developed. However, the problem of very small object tracking stays unsolved, in particular when are considered simultaneously moving camera and real time constraints.

So in this paper, we propose an original way to detect what we have called small objects and to track them in video sequences. Our approach consider the difficult case of both a moving camera and real time constraints. In order to detect the small objects, which are represented by some details in the images, we decide to tackle the opposite problem (detecting non-detail pixels) and removing the details from the image. It is then possible to consider only disappeared regions (small objects) and to track them along a video sequence.

We will first briefly present how the problem of small object tracking has been solved in the literature. Then we will describe our approach which consists of details removal using multiresolution filtering. A postprocessing is then necessary to cluster the different ”detail” pixels in ”small object” regions. The results obtained on every image can be merged to track the small objects along a video sequence. This procedure will be also presented. Finally some results will be given in order to illustrate our contribution.

2. The problem of small object detection and tracking in the literature

As it has previously been noted, small object tracking is a difficult problem. This can be justified by two main reasons. First, a small object is represented by a small set of pixels (typically between 10 and 100) in the image, and so does not contain enough information to be learned or modelled correctly. Secondly, due to their small sizes, small objects can often be mistaken for noise present in the image.

Among the few solutions given to solve the problem of small object tracking, we can consider a static camera which is the first and simplest case. It is then possible to extract moving regions by background or successive frame differencing. The difference image is further analysed to detect small objects. In the context of ball tracking, Ohno *et al* [3] detect in football games the ball by assuming it is a small moving region characterized by a color different from players, whereas Pingali *et al* [4] use in tennis games color to distinguish between the ball and players.

When dealing with a moving camera, the complexity of the problem is increased. Here we can cite the work from [1] and [6] which both focus on football game analysis. Both approaches are based on 3 steps: background substraction, ball detection, and ball tracking.

The last case considers simultaneously a moving camera and a real time framework. Sanderson *et al* [5] deal with boat tracking in sea images. Using the Fast Fourier Transform, it is possible to detect and remove the background area. The remaining regions are then analysed and tracked among the video sequence. However, this approach assumes that the background can be characterized by a frequentiel model, which is not always the case. So we propose another small object detection method, which is based on details removal from the image.

3. Details removal

As small objects consist of detail pixels, their detection can rely on a segmentation between details areas and other areas. We use a principle from [2], which considers a multiresolution framework to affirm that the coarser the resolution is, the less visible the details are. Details can then be removed by a multiresolution filtering.

From this result, we have now to locate these detail pixels and to cluster them into regions which correspond to small objects. So we compare the filtered image with the original image. For every couple of pixels, the euclidean distance is computed and thresholded, which allows to generate a binary image. This image is then analysed to estimate the position of small objects. So we scan the binary image with a shifting window to cover all possible positions with an overlapping property. For each window position, we determine the connected components and compute their size (*i.e.* the number of detail pixels they contain). If the size belongs to an interval defined according to the context of the application, we label the associated component as a candidate for a small object.

4. Tracking by temporal integration

The detection method presented previously gives for each frame some regions which are candidates to be considered as small objects, but also false candidates due to noise. Here we will show how to track small objects along a video sequence but also to eliminate these false candidates.

For every candidate region, we consider two different features which are its position and its shape. The tracking step consists then in the bidirectionnal matching of detected regions (using their features) between two successive frames. If a couple of regions has been matched in the two directions, we assume it is correct and we keep the real small object to track it in the following frames of the video sequences. Otherwise, it is assumed to be noise and it is discarded.

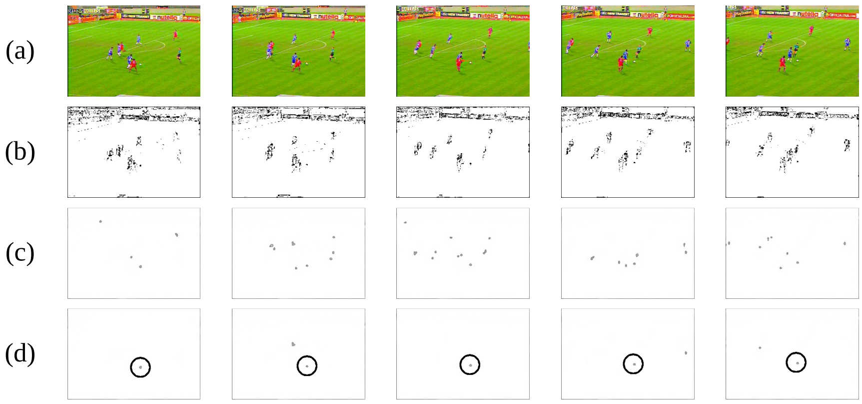
5. Results

In order to validate our contribution, we consider the problem of ball tracking in far shots, in the context of football game analysis. The figure 1 illustrates the different steps of our approach, which is able to process 15 images per second on a 1.7 GHz PC. In our application, size of small objects is assumed to be lower than 100 pixels. So we use the following parameters: a multiresolution pyramid depth equal to 4, a threshold equal to 60, a scanning window of size 50×50 pixels, and an object size interval equal to $[10, 70]$. In order to improve the clearness of the results shown in figure 1, we have surrounded manually the object with the highest duration in each frame. As we can see, the ball is correctly tracked whereas other detected objects which correspond to noise are discarded.

6. Conclusion

In this paper, we were dealing with tracking of very small objects in video sequences. This problem stayed unsolved when considering simultaneously a moving camera and a real time framework. Assuming small objects are represented

Figure 1. Tracking example on sample frames (left to right) of a football game video sequence: input image (a), details removal (b), candidates detection (c), and small objects tracking (d).



by detail pixels in the image, we proposed here a solution based on details removal using multiresolution filtering. A local analysis is then performed to detect candidate regions. A final interframe matching phase allows to isolate real small objects from noise and to track the objects along a video sequence.

The method has shown good performance and efficiency (about 15 Hz), even in the case of a moving camera. In order to process frame at real time (*i.e.* 25 Hz), we consider to implement the proposed method on a multiprocessor workstation. The use of other region features will also allow to improve the robustness of the method, which has to be validated on other kind of video sequences.

References

1. Y. Gong, L.T. Sin, C.H. Chuan, H. Zhang, and M. Sakauchi. Automatic parsing of tv soccer programs. In *IEEE ICMCS*, pages 167–174, 1995.
2. S. Lefèvre, L. Mercier, V. Tiberghien, and N. Vincent. Multiresolution color image segmentation applied to background extraction in outdoor images. In *IS&T CGIV*, pages 363–367, 2002.
3. Y. Ohno, J. Miura, and Y. Shirai. Tracking players and estimation of the 3d position of a ball in soccer games. In *IAPR ICPR*, pages 145–148, 2000.
4. G.S. Pingali, Y. Jean, and I. Carlbom. Real time tracking for enhanced tennis broadcasts. In *IEEE CVPR*, pages 260–265, 1998.
5. J. Sanderson, M. Teal, and T. Ellis. Characterisation of a complex maritime scene using fourier space analysis to identify small craft. In *IEE IPA*, pages 803–807, 1999.
6. Y. Seo, S. Choi, H. Kim, and K.S. Hong. Where are the ball and players ? Soccer game analysis with color-based tracking and image mosaick. In *ICIAP*, pages 196–203, 1997.