

ROTATION DETECTION USING THE CURL EQUATION

Daire Lennon, Naomi Harte & Anil Kokaram

Department of Electrical & Electronic Engineering,
Trinity College Dublin, Ireland
lennondh@tcd.ie, nharte@tcd.ie, anil.kokaram@tcd.ie

ABSTRACT

Rotational types of motion can often be seen in video sequences. However, not a lot of research has been done to investigate rotational motion models for use in video. Analysing this unique type of motion could be very useful. For example, if the centre of rotation of a spinning object can be efficiently identified, extraction and tracking of it can be made easier by grouping points moving at the same radial speed. It could also improve compression by recording rotation variables. In this paper, we introduce a method for finding the centre of rotation of a rotating object and a basic approach for modelling the rotation for improved image quality. The method requires an initial block based translational motion field.

Index Terms— rotation detection, rotation estimation, curl

1. INTRODUCTION

Motion Estimation is an essential part of most modern digital video encoding standards such as H.261 or MPEG. Frames are divided into blocks and the motion of each block is estimated from the previous frame. The motion of each block is represented by a single translational motion vector and a motion vector field is created for each frame. Therefore each frame can be considered a motion compensated version of the previous frame. In doing this, high compression rates can be achieved as only the original image and the motion vectors are required for reconstruction.

Although rotation is a well understood area, rotational modelling techniques are complex and their accuracy is still in question. Hence in video encoding translational motion estimation is the typical motion model used. This can produce poor results when reconstructing rotating objects, in particular at the object's edges, where blocking artefacts can become more apparent.

Rotational motion can often be seen in video sequences. The moving wheels of a car, skaters spinning on ice or a high diver performing somersaults are some examples of rotational motion that could be captured on video. If the centre of rotation of a spinning object can be efficiently identified, extraction and tracking of it can be made easier, as points of the object at the same distance from the centre move at the same speed, and can be grouped together easily.

Existing rotation analysis mainly focuses on transformation techniques. In [1] and [2], Fourier transform is applied to estimate the rotation angle. The wavelet transform is also used to estimate the motion of rotating objects in [3]. In [3] a new continuous wavelet transform was constructed that could be tuned to both translational and rotational motion. The parameters of analysis that are taken into

Thanks to Irish Research Council for Science, Engineering and Technology for funding this research, Grant No. RS/2005/68.

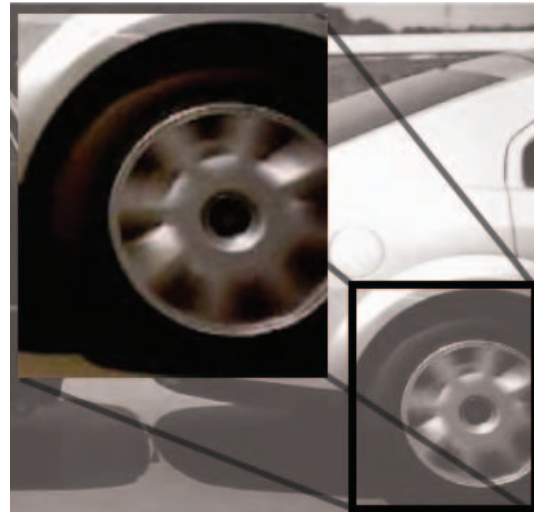


Fig. 1. An example of artefacts as seen on a rotating wheel. The fine detail is lost due to edge blurring in the motion estimation. There is a loss of still image quality.

account in the rotational wavelet transforms are space and time position, velocity, spatial scale, angular orientation and angular velocity. The analysis techniques introduced in these papers is often computationally expensive, and requires a well structured motion field that fits the mathematical models to work well.

The method of rotation detection developed in this paper has already been used in a Psychology project [4], where the rotating object in question was the top of a child's head. The particular event of interest was rotation of the child's head and that information was used to finely parse a video sequence. These further segmented sequences were to be automatically analysed and particular motions extracted.

While [4] is concerned with a particular kind of event, this paper introduces a technique to improve rotation modelling in an arbitrary image sequence. In Section 2, a novel method for rotation detection and centre of rotation identification is introduced. This method uses a translational motion field from a block based motion estimator. In Section 3 a means to segment the rotating area is shown. In Section 4 a method for estimating the amount of rotation occurring is introduced. Also in this section a method is developed to rotate the rotating section of the image instead of using the translational motion field to reconstruct this part of the image. The experimental setup for evaluating the efficiency and effectiveness of this method

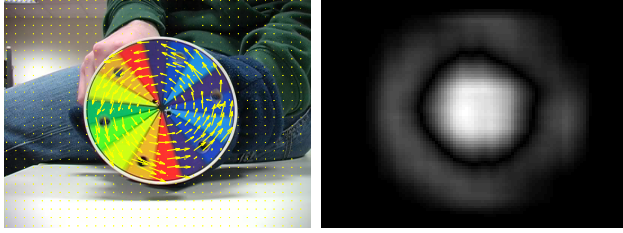


Fig. 2. An example sequence of a rotating rainbow disc. The number of motion vectors overlaid have been reduced for clarity. The second image is the resulting curl measurement for the entire motion field.

are shown in Section 5. This is followed by a summary in Section 6.

2. ROTATION DETECTION

By measuring the curl of the motion field it is possible to detect rotation. The motion field is initialised by translational motion field on a block basis. Gradient Based motion techniques introduced in [5] are used here. Curl is defined by the equation below.

$$C(d_1, d_2, \mathbf{X}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{d}{dx} & \frac{d}{dy} & 0 \\ d_1 & d_2 & 0 \end{vmatrix} = \vec{k} \left(\frac{d(d_1)}{dy} + \frac{d(d_2)}{dx} \right) \quad (1)$$

$\frac{d}{dx}$ and $\frac{d}{dy}$ are horizontal and vertical gradients respectively. \vec{i} , \vec{j} and \vec{k} are the orthonormal basis for the vector space. d_1 and d_2 are the estimated translational motion vector components in the x and y directions at pixel \mathbf{X} . The curl is essentially the sum of the derivative of the x component in the y direction with the derivative of the y component in the x direction.

The curl measurement at each pixel is smoothed using a Gaussian filter of suitable size for the given resolution. The curl measurement gives clear evidence of rotation in the form of a large peak at the centre of the rotation. The centre of rotation is taken to the peak maxima. An example is shown in figure 2. Clockwise pointing motion vectors will give a positive central peak value while anti-clockwise motion vectors will give a negative central peak value. This is useful for estimating the angle of rotation later on, as it tells you that there is only one direction to search in.

During a rotation event certain observations can be made about the curl peak that correlate with rotation. The value of the maxima can be directly correlated with the rate of rotation and the area the peak covers is directly related to the visible area of the rotating object. The location of the main rotational peak can be considered constant during rotation and measuring the fluctuation in position helps indicate whether a peak is a rotational peak or not. These correlations can be seen in figure 3. The data was observed for an example sequence of the rotating disc in figure 1.

As can be seen in figure 3 whenever there is a rotation there is a rise in the curl maxima value and a fall in the corresponding distance values from curl maximum to curl maximum from frame to frame. The fall in distance value indicates a increase in curl position stability which is expected during rotation.

The next step in replacing a translational motion model with a rotational model is to find the area of the rotating object. Then an

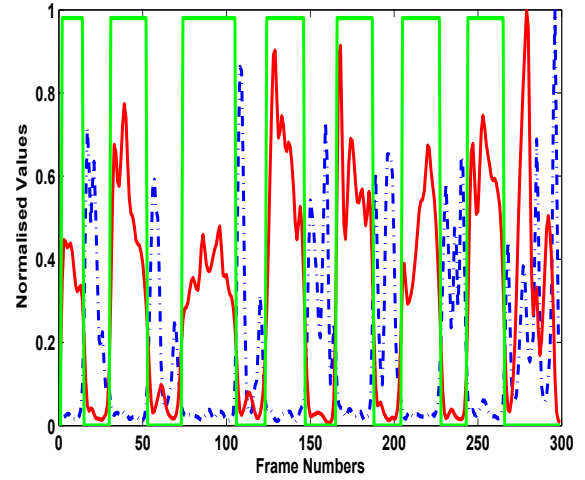


Fig. 3. In the graph above the red solid line plot represents a sample normalised curl maxima values. The dashed blue line plot above is the normalised distance values from curl maximum to curl maximum for consecutive frame. The green vertical line plot is the manually noted beginning and end points of rotation in the video sequence.

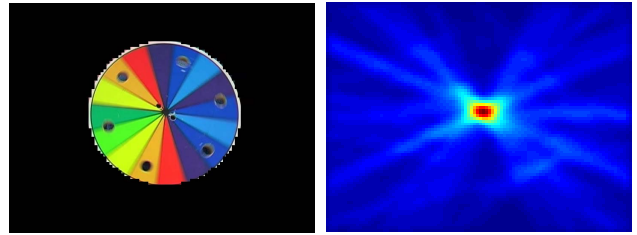


Fig. 4. An right image is an example Rotation image with overlaid estimated rotation mask. The left image is the accumulation of the lines perpendicular to the motion vector for a frame.

estimate of the centre of rotation and angle of rotation is needed. The approach presented in this paper is relatively simple one which can be implemented with relative ease.

3. ROTATION SEGMENTATION

For a more accurate representation of rotation, a robust mask covering the rotating region is required. This was done by exploiting the proximity of the perpendicular motion vector lines to the centre of rotation, an example is shown in figure 4. This is based on the principle that perpendicular lines to tangents (the motion vectors) of a circle (the rotating object) always pass through the centre of rotation. All perpendicular lines that pass within a reasonable tolerance of the centre of rotation can be considered part of the rotating object.

There is however a problem with short motion vectors that can cause large deviations from what is expected. To reduce this effect the vectors with a length shorter than a threshold are ignored. This improves the overall accuracy of the mask estimate.

There is a need to remove any random motion vectors outside the

object that happen to pass close to the centre of rotation by chance. To do this consider the distance from the centre of rotation to each motion vector position. A histogram of distances from each motion vector position to the centre of rotation is used to develop a threshold radius, inside which is the rotating object.

Given the observed Gaussian nature to the radius histogram, the threshold radius R was set to be the upper bound of the 95% confidence interval.

$$\text{Threshold Radius } R = M + 1.96 \times SD$$

Where M is the histogram mean and SD is the histogram standard deviation. The segmentation mask was then set to be circular with the centre located at the curl maxima and radius equal to R , as shown in figure 4. Any pixels within the radius R are set to one and those outside are set to zero. When this is multiplied by the image only the rotating object is visible, an example of which can be seen in figure 4.

4. ROTATION ESTIMATION

The model for pure image rotation is as shown in equation 2. The model is only applied to the masked part of the image found in the previous section.

$$I_n(h, k) = I_{n-1} \left(\begin{bmatrix} h \\ k \end{bmatrix} \times \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \right) + e \quad (2)$$

In equation 2, h and k are the pixel positions in the frame matrix measured with reference to centre of rotation. θ is the angle of rotation and e is the error difference between the two frames. From equation 2, the error function, used in the simplex method, is derived. The Simplex method finds the smallest value of θ required to minimise the mean square error E of the error function 3 below.

$$E = \sum_{h,k} [I_n(h, k) - I_{n-1}(h\cos\theta - k\sin\theta; h\sin\theta + k\cos\theta)]^2 \quad (3)$$

The Simplex search technique of [6] is used to minimise the error function E . This is a direct search method that does not use numerical or analytic gradients. If n is the length of x , a simplex in n -dimensional space is characterized by the $n + 1$ distinct vectors that are its vertices. In two-space, a simplex is a triangle; in three-space, it is a pyramid. At each step of the search, a new point in or near the current simplex is generated. The function value at the new point is compared with the function's values at the vertices of the simplex and, usually, one of the vertices is replaced by the new point, giving a new simplex. This step is repeated until the diameter of the simplex is less than the specified tolerance.

Once the θ has been calculated the masked image section of frame $n - 1$ is rotated θ radians around the previously determined centre of rotation. The mask itself is transformed to align with the position of the rotated object from frame n . This mask can now be used to switch smoothly between translational and rotational motion compensation.

5. EXPERIMENTAL RESULTS

In figure 5 a sequence of frames are shown overlaid with the forward pointing motion vectors, these estimate the position of each block in the next frame. The second row in figure 5 show the estimated

reconstruction of the next frame using the motion vectors and the previous frame, as it can be seen this is less than ideal as there are many blocking artefacts and distortions. The third row show the rotation masking for the previous frames. The bottom row of figure 5 are frames constructed using the rotational motion model discussed in this paper. It can be seen that it gives a much more accurate representation of the image with relatively minimal edge distortions. Over 10 frames of the sequence, the prediction error of the mask and rotation estimate was compared to the translational motion model. A reduction of 20% was observed.

6. CONCLUDING REMARKS

This approach worked well when detecting rotation and the curl equation can be used to find some very interesting features about rotation. There are however some limitations to this approach. First of all, the method works well when the the rotating object is parallel to the image plane of the camera, but in many other instances of rotation in real scenes this is not true. Transforming the vector field in the region of rotation centres should help to segment rotating objects that are not exactly parallel and this is currently being considered.

In addition, the rotating object must have sufficient texture and must be rotating slowly enough to be detected by the motion estimator. If these conditions are not met then the rotation would not be obvious to the human eye anyway, and therefore finding rotation would have no benefit. Future work will include the possibility of using this approach directly on MPEG stream motion vectors. The goal would be to improve the motion estimation and compression ratio. If this approach could be included in MPEG preprocessing, then circular motion could be more accurately represented and visible blocking artefacts reduced.

7. REFERENCES

- [1] S. Alliney, "Digital analysis of rotated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 5, pp. 499–504, 1993.
- [2] Hiroyuki Onishi and Hisashi Suzuki, "Detection of rotation and parallel translation using hough and fourier transforms," in *IEEE International Conference on Image Processing*, September 1996, pp. 827 – 830.
- [3] Mingqi Kong, Jean-Pierre Leduc, Bijoy Ghosh, Jonathan Corbett, and Mladen Victor Wickerhauser, "Wavelet based analysis of rotational motion in digital image sequences," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 1998, pp. 2777–2780.
- [4] Daire Lennon, Naomi Harte, Anil Kokaram, Erika Doyle, and Ray Fuller, "A hmm framework for motion based parsing for video from observational psychology," in *IEEE Irish Machine Vision and Image Processing Conference*, September 2006.
- [5] A. C. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, Springer Verlag, ISBN 3-540-76040-7, 1998.
- [6] J.C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder-mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

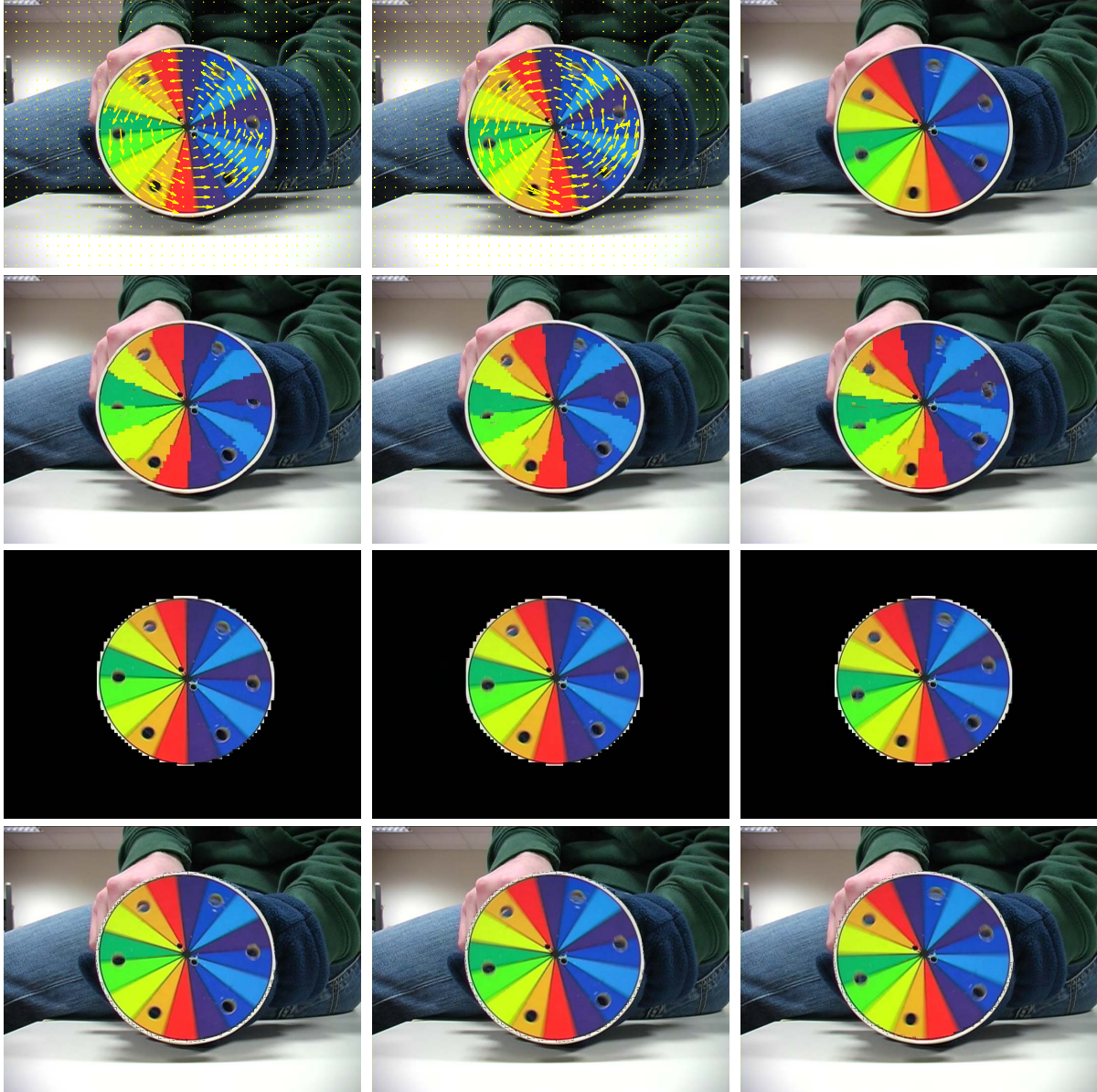


Fig. 5. The top row of images show frames 36 to 38 overlaid with translational motion vectors. The second row of images show the motion compensated frame $I_{n-1}(x + d_{n,n-1})$, where $d_{n,n-1}$ is the translational motion for a block in frame $n - 1$. The third row of images are the masked portions of rotation for frame 35-37. The bottom row of images show the reconstructed frames 36-38 from frames 35-37 using the rotational model.