Inferring Changes in Intrinsic Parameters From Motion Blur

Alastair Barber[a,b,*], Matthew Brown[a], Paul Hogbin[b], Darren Cosker[a]

[a] *Centre for Digital Entertainment, Computer Science, University of Bath, Bath, UK*
[b] *Double Negative Visual Effects, 160 Great Portland St., London, UK*

Abstract

Estimating changes in camera parameters, such as motion, focal length and exposure time over a single frame or sequence of frames is an integral part of many computer vision applications. Rapid changes in these parameters often cause motion blur to be present in an image, which can make traditional methods of feature identification and tracking difficult. In this work we describe a method for tracking changes in two camera intrinsic parameters - *shutter angle* and scale changes brought about by changes in *focal length*. We also provide a method for estimating the expected accuracy of the results obtained using these methods and evaluate how the technique performs on images with a low depth of field, and therefore likely to contain blur other than that brought about by motion.

1. Introduction

Estimating motion of a camera system, both in terms of *extrinsic* (camera movement relative to the world coordinate system) and *intrinsic* camera changes (such as changes in focal length) is an important aspect of many computer vision applications. Accurate estimation of these changes throughout a film sequence is an essential part of the Visual Effects (VFX) process, as without this information, computer generated assets, such as characters, scenery and effects, cannot be applied convincingly to live-action footage. Often, in order to determine changes in the camera parameters, it is necessary to track individual feature points over two or more frames after filming has taken place, or use additional camera mounted hardware such as a motion capture rig, inertial measurement devices, and other devices for tracking physical changes to the lens parameters. Commonly, the process of determining changes in camera parameters after filming is referred to as *match-moving*. This is a process that uses structure-from-motion computer vision techniques to estimate both camera motion and 3D scene structure using corresponding feature points over multiple frames [13, p. 207]. This process can often be time-consuming, and require the input of a skilled operator in order to produce an accurate camera track from even automatically detected and matched feature points. In the case of using additional hardware, this presents challenges such as gaining acceptance on set for installation, and the additional expense of equipment and operation. There are also often many situations where such equipment would be impractical - such as outdoors or at sea, due to the reliance on additional infrastructure. However, recent developments in electromechanical sensors has allowed for the manufacture of gyroscopes and accelerometers that are both low cost and small. These devices are now starting to be included within cameras and can easily be mounted to them in order to provide information about their motion during filming. Examples of applications of such camera mounted devices range from assisting determining scene geometry [11] to correcting for distortions introduced by motion and camera rolling shutter [5]. One of the most significant challenges with using inertial measurement sensors to measure motion of the camera is that only changes in acceleration or rotational velocity are recorded. This can lead to significant errors in determining absolute position by integrating this data [12], and as such are rarely suitable for tracking camera motion when used alone. Devices which track physical changes in lens parameters are now commonly used in production environments and have gained acceptance across the industry - however they must be accurately synchronised to the video captured by the camera. Whilst this is now a quick process, occasionally it may not be completed correctly (if at all) for each shot, and manual alignment of the data in post-production is a time consuming and hence expensive task.

Accurate feature tracking is a reliable method of determining accurate camera motion estimations, and is an active area of research. However, there are several cases where it is difficult to get an accurate track, most noticeably when there is a fast unpredictable motion of the camera, which also often leads to a considerable amount of motion blur being present in a frame, making features undetectable. Another common method for determining camera movements is to make use of a method known as 'Optical Flow' across an image. In this process, a dense correspondence for each pixel across two frames is calcu-

*Corresponding Author
Email addresses: a.e.barber@bath.ac.uk (Alastair Barber), m.brown@bath.ac.uk (Matthew Brown), hogbin@dneg.com (Paul Hogbin), dpc@cs.bath.ac.uk (Darren Cosker)

lated. Assuming that there are a sufficient number of stationary objects in the scene, the camera's movement can be calculated using this correspondence information. Similarly to automatic feature detection and matching, the process of calculating the optical flow across frames also suffers from degradation in the presence of large quantities of motion blur.

In [17], the authors present a method for determining dense optical flow in the presence of spatially-varying motion blur. This method produces good results, however calculating optical flow over an entire image can be a computationally expensive process. In [6], the authors present a method of determining in real-time and using a single motion-blurred frame, an estimate for camera rotation - using characteristics of the motion blur directly, and without selecting or matching any features from the image.

In our previous work [1], we used motion blur induced onto an image by changes in focal length and camera rotation to track changes in two camera intrinsic parameters - namely focal length and shutter angle. We used accurate hardware tracking of changes in camera parameters (the focal length change of a lens and camera rotation) to gather ground truth datasets and validate our algorithms. We also demonstrated how, in a situation where unsynchronised data from certain sensors was available alongside blurred footage, the blur patterns from frames in this footage could be used to accurately synchronise the external data with camera frames. One of the main limitations of the approach presented in [1] is that in order for an accurate estimate of focal length to be produced, there must be a sufficient amount of motion-induced blur present in the frame, along with sufficient visual texture (in this case, sharp edges). In the following sections, we give an expanded description of our method as presented in [1] for determining shutter angle and scale change brought about by focal length change. In addition to this, we present an extension to this method for validating the accuracy of such results across two new datasets in differing conditions. We also investigate the effects of a shallow depth-of field (and hence images likely to contain a significant amount of blur irrespective of motion) on both our method.

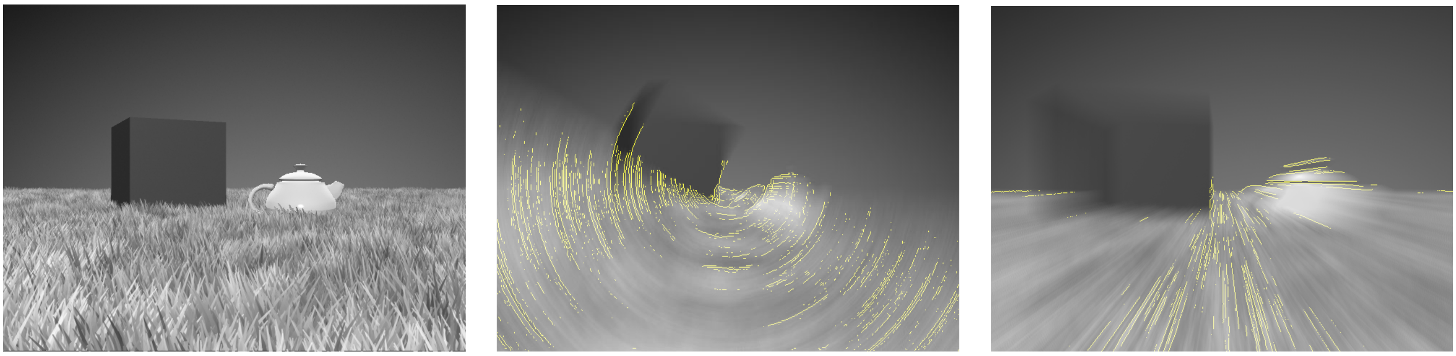
2. Background

Our main motivation for this work is to improve the process of 'Matchmoving' for use in Visual Effects. In particular, we are interested in accurately estimating changes in camera parameters automatically and from scenes that would cause traditional structure from motion techniques based upon feature detection and matching to fail. Motion blur is often present in footage, and it is not uncommon for it to be considered a desirable artistic effect by directors in order to convey a sense of fast movement to the viewer [4]. This can often present challenges in determining an accurate camera track [13, pp140-143], as many current techniques for feature identification and matching rely on there being sharp corners or changes in image intensity being visible. Motion blur severely reduces the occurrence of these in an image. However, recent work has looked at using the characteristics of induced motion blur alone to determine parameters of a scene in order to avoid this limitation.

Using Motion blur directly to determine parameters of a scene is an area of current computer vision research. [9] presents a method of determining speed of a moving vehicle from a blurred image, whilst then using this information to de-blur the resulting image. Other methods, such as the one presented by Rekleitis [14] use the direction and magnitude of motion blur in the process of estimating optical flow in an image. Later work, in [17], parameterises each frame as a function of both pixel movement and motion-blur. In [17], the authors determine the derivative of the blurred frame with respect to both the motion and the blur, where the blur itself is a function of motion. Furthermore, if the exposure time is known as a fraction of the frame (*shutter angle*), the result can be further optimised. Recent work in [7] makes use of data captured from a 3D pose and position tracker attached to the camera to aid in the calculation of optical flow in images affected by motion blur. As the level of motion blur in an image is typically directly related to the exposure time of the frame, [10] and [16] use a method with a hybrid camera capturing both high and low frame-rate images of the same scene to correct images exhibiting motion blur.

Presented by Klein and Drummond in [6] is a method for determining the rotation of a camera during a single-frame exposure resulting in motion blur. In this work, the axis of rotation is derived by selecting a point through which the most normals to the edgels at a set of 'edgel' (points along an edge) points coincide. This algorithm builds on the observation that areas of motion blur will typically form edges in the image. Figure 1 shows a synthetic animation that has undergone motion blur whilst the virtual camera has been rotated, and the results of this image having undergone Canny edge detection.

In the case of the scene in figure 1, the algorithm described in [6] will estimate the centre of rotation to be at the centre of the image plane - the Z axis. In order to handle rotations around the X and Y axis, the normal line to the edge at each edgel site is expressed as the intersection of the image plane with a plane passing through the origin and and edgel site. Once the centre for rotation has been accurately determined using RANSAC (and optimised using a Levenberg-Marquardt based algorithm), the magnitude of rotation can be determined from analysing the blur along its direction, with the intensity of pixels in the image being sampled in concentric circles centred at the estimated axis of rotation. In [6], rotation magnitude is estimated under the assumption that the blur length cannot exceed the shortest intensity ramp produced by an intensity step in the scene (i.e., the least blurred feature). Under the further assumption that the largest intensity step in each scene will span approximately the same in-



Original image | Image with motion blur from Rotation | Image with motion blur from change in Focal Length

Figure 1: Images Blurred from Camera Rotation and Focal Length Changes with Resulting Canny Edge Detection

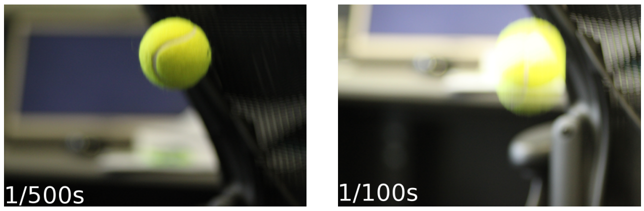
tensity increase, the gradient of the steepest ramp to span this increase will therefore be inversely proportional to the length of the motion blur, and thus the magnitude of rotation from the camera. Their work highlights a number of important limitations in using motion blur to determine changes in camera parameters, most notably that from a single frame alone, it is not possible to determine the direction (or sign) of rotation. For this reason, it is only possible to compare the results of this algorithm with normalised values of rotation from a rate-gyroscope or other method for determining ground truth.

2.1. Intrinsic Parameters

The intrinsic parameters we consider in this work are *focal length* and *shutter angle.*

If the focal length of a lens were to change whilst the sensor or film is exposed, it could be expected that the image will experience motion blur in a similar fashion to those described in the previous section due to changes in the field of view. An example of such an image is also shown in Fig. 1. Although the entire image has been scaled by a single value, it is apparent that different parts of the image are blurred by differing amounts, specifically - towards the centre of the image edges will still appear sharper, despite being scaled, than towards the outside. It is also clear that the 'edges' introduced by this blur converge towards the centre of the image, in a similar fashion to a translation of the camera originating from the centre of the image.

When a frame is captured, the image sensor, or film, is exposed for a short amount of time. Often, this amount of time is known and controlled by the camera operator - however there are occasions where this would be an unknown value, such as in cameras with an automatically controlled exposure. Fig. 2 shows two extracts from two video sequences of a ball falling under gravity. The left hand panel is a frame from a sequence shot with an exposure time of 1/500th of a second, whilst the right hand panel shows a similar scene captured with an exposure time of 1/100th of a second. In both frames, the ball falls at an identical speed, and in both cases the *frame rate* was set to 25 frames per second. Therefore, the left frame would be exposed for $\frac{1}{500} \div \frac{1}{25} = 0.05$ of the frame time and the right hand frame for $\frac{1}{100} \div \frac{1}{25} = 0.25$. It can be seen from Fig. 2, the frame with the longer exposure time as a fraction of the frame exhibits the largest amount of motion blur. Historically, this fraction of time for which the frame is exposed is determined by the *shutter angle.* This is so called as in cameras with mechanical shutters consisting of a rotating disk with an adjustable sector with which to expose the film, the shutter angle referred to the angle of opening of this sector. In the example from Fig. 2, the shutter angle of the second frame would be $360° \times 0.25 = 90°$, and a frame for which the exposure time is half the frame time would be $180°$. Throughout this work, for simplicity, we refer to the values for shutter angles as fractions of the frame time.



Shutter Angle 18° | Shutter Angle 90°

Figure 2: Illustration of Shutter Angle and Motion Blur (25fps)

3. Method

3.1. Measuring Focal Length Change from a Single Frame

In the case of a single motion-blurred frame undergoing rotation, we use Klein and Drummond's original method to calculate the rotation, R around a 3D axis for that frame. In our work, we focus on scale change in the 2D image coordinate system. We also extend this method to determine a scale change brought about by a change in focal length without other motion. In our work, we focus on a scale change taking place in the 2D image plane, with the principal point of the lens being at the centre of the image.

As shown in Fig. 1, the change in focal length (assuming the camera is not rotating or translating) adds motion blur to the image in a fashion similar to a translation towards the principal point of the image plane. Unlike the method used by Klein & Drummond to estimate for rotation, there is no need to determine the centre of the transformation as we can assume that the direction of the blur will always be towards the principal point of the image plane. Therefore, in order to determine the magnitude of blur, the intensity I of the image along several radial lines L, is sampled from the edge of the image inwards (Fig 3). The number of radial lines depends on the size of the image, and are sampled starting at locations on the edges of the image spaced 10 pixels apart. Therefore, for a 640×480 image, there would be $2 \times 64 + 2 \times 48$ lines sampled. This profile is then searched for the first occurrence of an intensity step change greater than a threshold value - and the length of this change (and image position of the start and end) is recorded. In a similar fasion to the authors of [6], we choose a threshold value in order to avoid under-estimating the length of the blur, and only consider ramps which span a large intensity change (over 50 grayscale levels) in order to detect large isolated intensity steps (representing edges) in the image. The first occurrence of the step-change is selected because edges are expected to be less blurred towards the centre of the image, and hence the shortest intensity ramp will always correspond to a minimally blurred edge towards the centre of the centre of the image. Unless the scale change is very large, the likelihood is that this edge towards the centre of the image will not have been affected by the scale change or motion blur, and will therefore represent a scale change of zero, regardless of the true change in scale. As the origin of the scale change will be the centre of the image Eqn.1 describes this relationship between an image point u and the point u' after a change in focal length f: Δf.

$$\begin{aligned} u &= f\frac{X}{Z} \\ u' &= (f+\Delta f)\frac{X}{Z} \\ \frac{u'}{u} &= 1 + \frac{\Delta f}{f} \end{aligned} \tag{1}$$

Where X is a scene point of distance Z from the front nodal point of a lens.

Figure 3 shows the location of a blur region as detected by this algorithm in a synthetically blurred image, and Fig. 4 the locations of all blur regions over the image.

After a pair of points has been obtained for each radial line, a RANSAC based algorithm is used in order to determine the geometric transformation between the sets of points. In this process, the start and end points of the maximum gradient ramps from the radial search lines are

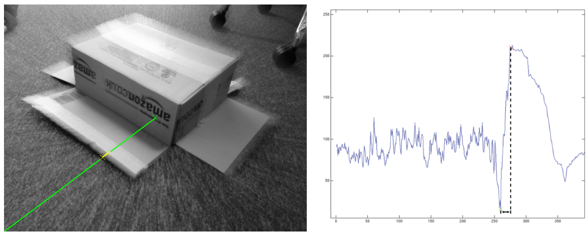


Figure 3: A line sample location (left) and profile (right). The peak gradient has been highlighted and location marked on the image.

represented as their respective image coordinates. The geometric transform brought about by a change in scale is then estimated to produce an estimate of the scale transform, using the points identified at each radial line. To achieve this, we adapt the standard RANSAC algorithm to take into account the observation that measuring the magnitude of motion blur by searching for the maximal gradient ramp will always produce an overestimate for the blur magnitude. This would be because even in the case where there is no blur, the sharpest edge might be several pixels in extent, and in practice, in an image with moderate motion blur, will extend several pixels beyond the blurred region. Because of this, the error metric used in the RANSAC based geometric estimation is weighted to apply a higher penalty to estimations that produce an under-estimate of the scale magnitude. This is done by changing the model of our system in order to achieve a result that match with the assumption that measuring the length of a blurred edge will result in an over-estimate of the true scale change.

In this process, instead of finding a hypothesis to maximise the number of start and end points for blur that comply with $((r' - r)^2 < \epsilon^2)$ where r' and r are the measured and predicted radial displacements, we maximise $\sum((r' - (r + \epsilon))^2 < \epsilon^2)$. By using this method, in order to be considered an inlier, r' must be in the range r to $r + 2\epsilon$, as opposed to $r - \epsilon < r' < r + \epsilon$ as in a traditional RANSAC procedure. The upper limit of this range: $r + 2\epsilon$ was chosen as a limit arbitrarily and produces good results, however it should be noted that other values, or the use of methods such as Least Median Square estimate, or MLESAC could be used to determine this value, although these are not evaluated in this work. This method provides an accurate estimate of the transformation between the points - whilst also rejecting outliers in the sets of points.

As described in Section 2.1, the shutter of the camera will only be open for a fraction of the frame time depending on the *shutter angle*. The estimate for scale change from motion blur will only take into account the time for which the shutter was open, and not the overall frame.

3.2. Measuring Rotation Between Two Frames

The optical flow of two motion-blurred images can be calculated using the baseline method described in [17].

Figure 4: Blur length estimation along all radial lines

Then, a set of feature points in the first frame are sampled using [15], and their flow vectors used to calculate corresponding points. As it is expected that there will be some outliers, we use a RANSAC algorithm similar to that described in Klein & Drummond to determine a consensus set of matching points, in order to determine rotation. Assuming a correct pair of point matches, $\hat{p}_1$ and $\hat{p}_2$, where $\hat{p} = [x, y, 1]^T$ is a homogeneous point in the image coordinate system, the line joining these points will be described as $L_p = \frac{\hat{p}_1 \times \hat{p}_2}{|\hat{p}_1 \times \hat{p}_2|}$. As $\hat{p}_1$ and $\hat{p}_2$ are homogeneous coordinates, the line $L = (a, b, c)^T$ for which a point $\hat{p} = (x, y, z)$ lies on is specified by the equation $ax + by + cz = 0$. Assuming a further pair of correct point matches is available, and the normal line to these can be calculated, the point of intersection of these two normal lines (L_1 and L_2) should then be the centre of rotation. This is where using the homogeneous coordinate system is useful, as if the camera is rotating around a point not in the image plane (for example, its x or y) axes, the centre of rotation can still be represented in the image coordinate system, as the two normal lines from point estimates would cross at infinity, a point which can be represented in homogeneous image coordinates as $\hat{p} = (x, y, 0)^T$.

Candidate point pairs and the best estimate for rotation are selected using RANSAC. In this process, a pair of candidate points and their matches are selected, and the centre for rotation, C is calculated based on the method described above. The connecting line for every other point match is calculated, and the normal at the midpoint to this line L_N, along with the line L_C from this midpoint to the centre estimate, is calculated for each point pair. This is illustrated in Fig. 5. The angle between the line L_N and L_C, θ, is calculated for each point pair - and capped at a threshold value ϵ. In this work the value for ϵ is small, at 5 degrees, however should be varied by the user depending on the amount of candidate points expected (which can depend on the visual texture of a scene) and expected rotation magnitude.

The centre estimate producing the lowest sum of these angles is then selected as the rotation centre. This point

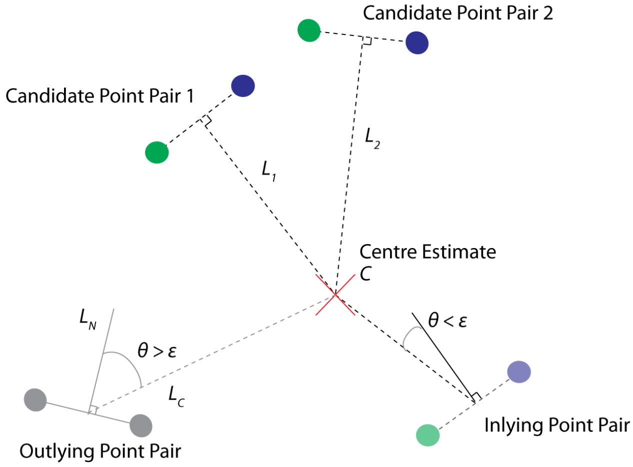


Figure 5: Illustration of Estimating the Centre of Rotation from Point Match Pair Candidates

is then normalised, and it's coordinates $C = (x, y, z)^T$ treated as a 3D point. The Least Mean Squared value for the angle between this point and the centre points between inlying point match pairs is then treated as the frame-to-frame rotation magnitude. Results obtained using this method alongside Klein and Drummond's single frame method - using synthetic and real image sets are shown in the following sections.

3.3. Determining Shutter Angle

By combining the results for rotation obtained from a single frame, and those from a pair of frames - it should be possible to calculate the exposure time of the frame as a fraction of the framerate, simply by dividing the motion magnitude obtained from blur by that of the frame-to-frame track. This calculation could further be simplified by using just the geometric distance between points identified by searching along the radial or circular profiles. However, it is envisaged that by performing the extra stages of rotation estimation will provide a more robust estimation for shutter angle. This is because both methods for determining rotation include the rejection of outliers as an important stage in the calculation of the magnitude.

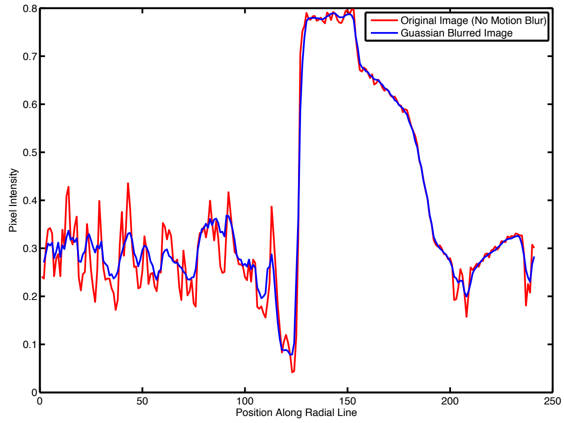
3.4. Determining Amount of Blur in an Image

It is envisaged that the methods presented previously will only work well if there is a sufficient amount of blur from motion present in the image. This is a limitation also highlighted by the authors of [6]. In order to evaluate the effectiveness of the method for accurately determining the scale change of different magnitudes across different sets of images, we propose a method for quantifying the amount of blur present across the whole image. Furthermore, it is proposed that this accuracy measure could be used to correct estimates over further footage of the same scene, given a ground truth for some initial data. This could be useful in such a situation where, for example, external hardware was being used to record the change in lens barrel and hence focal length position - and this hardware becomes unsynchronised or uncalibrated throughout the shot. Such situations are not uncommon and can require a large amount of work post-production to rectify. We would also typically expect the methods described here to be applied on a sequence of frames, some of which will not contain any change in focal length or rotation. As part of the process for estimating shutter angle from rotation (a change in an extrinsic parameter), it is possible to accurately deduce cases for which rotation and hence blur is zero using the optical flow method (results of which are shown by Fig. 8) which must be performed on each pair of frames. As previously stated, it is not possible to identify an blurred edge of length zero, so in the case of zero focal length change - the proposed algorithm will always return a result greater than zero. Classifying the blur characteristics of a frame with zero scale change would therefore allow for automatic identification of these frames

In the case of focal length from a single frame the following method is used to determine the amount of blur present in an image. We define blur energy ratio r_{blur} in an image as the average ratio between the energy of a profile of pixel intensities along a set of radial lines across an image, and the average energy of the same set of sample lines of the same image after having undergone a gaussian blur operation. In this work we used a Gaussian kernel $\omega = \left[\frac{1}{4} - \frac{a}{2}, \frac{1}{4}, a, \frac{1}{4}, \frac{1}{4} - \frac{a}{2}\right]$ where $a = 0.375$, and in order to produce a more significant result for the difference in energies across a radial profile, the difference between the top and the 3rd level of the Gaussian reduction pyramid is sampled. Similarly to the method used for determining scale change from motion blur, radial lines are sampled from the outside edges of the image inward - initialised at 10 pixel intervals along the edges of the image. The reasoning behind this is that an image that contains motion blur will have a lower energy (lower frequency of changes in intensity) than a sharp, non-motion blurred image - as described in earlier sections. However the ratio of energy between this motion blurred image and its gaussian blurred equivalent should be larger than the ratio of profile energy between a non motion-blurred image and its blurred equivalent. This is illustrated in Figure 6 and Figure 7, where it can be seen that for a non-motion blurred original image, there is a much higher frequency (and hence greater energy) of intensity change for the original image than the gaussian-blurred equivalent image. For the profiles shown in Figure 7, the frequency of changes in intensity for the original image is much closer to that of the Gaussian-blurred equivalent. We define energy as the sum of squared values of image intensity along the profile line, and sample along multiple profile lines, taking the mean ratio of energies across all lines over the image pair to be value for the difference in image energy.

Original Image with No Motion Blur (Sharp image)

Gaussian Blur of Sharp Image

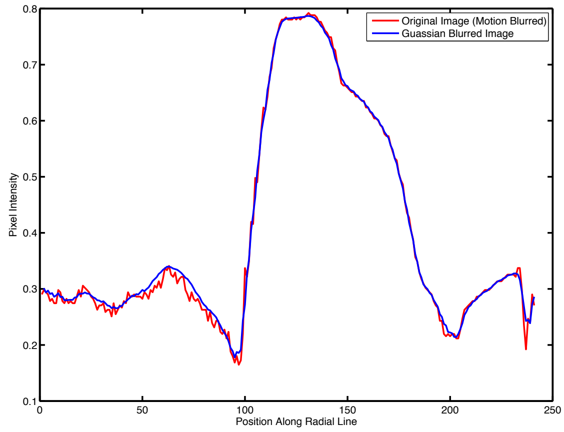


Pixel Intensity Profiles for Non-Motion Blurred Image and Corresponding Gaussian Image

Figure 6: Non-Blurred image and Gaussian Blurred Image with Corresponding Profile Lines

Original Image with Motion Blur

Gaussian Blur of Motion Blurred Image



Pixel Intensity Profiles for Motion Blurred Image and Corresponding Gaussian Image

Figure 7: Motion-Blurred image and Gaussian Blurred Image with Corresponding Profile Lines

4. Results

Presented in this section are the results obtained from a variety of tests, both on synthetic and real footage. In the case of synthetic images, a single static photograph had an animated scale change applied using the Nuke compositing tool (a 2D image manipulation package well suited to applying transforms, filters and animation and used widely in the post production industry). Motion blur for this set of images was then simulated for the specified shutter opening time at each frame.

Initially results are shown as in [1] for the raw output produced from running the algorithms for estimating changes in intrinsic values on a sequence of frames without first considering the amount of blur present in each frame of the sequence using the method described in Sec. 3.4.

For real image sequences, an external electro-mechanical *zoom encoder* was attached to the lens on the camera used to capture the footage. This is a proprietary device that uses a geared rotary encoder meshed with the zoom ring on the lens barrel to track change in rotational position of the ring. After a simple calibration and synchronisation, this data can be used to infer the focal length at a particular frame, independently from the image captured by the camera. Such devices are commonly used throughout the visual effects and post-production process as they provide a reliable method of measuring changes in camera parameters.

For the production of ground-truth values for camera rotation, the camera was rigidly attached to a high-end rate-gyro capable of determining rotation up to a speed $175°/sec$ with a standard error of $0.0005°/sec/\sqrt{\text{Hz}}$. [12] presents a comprehensive description of the specifications and sources of error in inertial measurement systems.

The values obtained from both the ground truth and original estimates of a real data-set for change in focal length are then used to calculate the expected error factor for each range of blur magnitude present in the frame. The ground truth magnitude for scale change is also used to validate that our measurement of blur present in a frame is effective. These error metrics are then used to attempt to produce a more accurate estimate of scale change from blur, using new footage of the same scene.

4.1. Synthetic Tests

To test the algorithms against a synthetic and known ground truth for a change in focal length, shutter angle, and rotation, the Nuke compositing tool was used to create an animated series of frames from a single image.

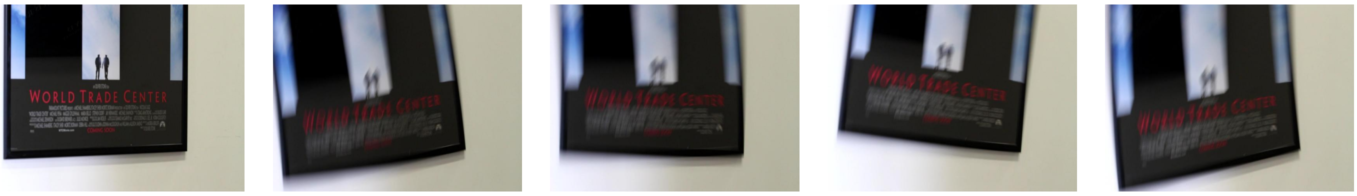
4.1.1. Focal length change from a Single Frame

Results for the motion estimates for a set of rotation changes and changes in focal length are shown here. In both cases, as it is not possible to determine the direction of motion from a single frame, all of the values for both focal length change and rotation are absolute values. Fig. 10 shows a plot for results obtained for determining the change in scale induced by a change in focal length. In panel *(i)*, the dashed blue and red lines should ideally be identical, and in the scatter chart in panel *(ii)*, the points should lie in an $x = y$ line. In this result, the chart in panel *(i)* also shows the change in scale corrected for the known shutter exposure time of the virtual camera, which should equal the frame to frame estimate of scale (the true scale in this case). For most frames, it can be seen that the raw estimation from blur overestimates the true scale value. This is to be expected, as if there is zero blur, the sharpest edge in the blur profile to be found (as described in Sec. 3) will still be at least one pixel (in practice on real photographs, this will likely be more) - which will therefore always result in some scale change being estimated. This is the effect that we aim to compensate for using the blur-information obtained using the method described in Sec. 3.4 to estimate the expected error of results of a scene, and the results for this when applied to a real scene are shown in Sec. 4.3.

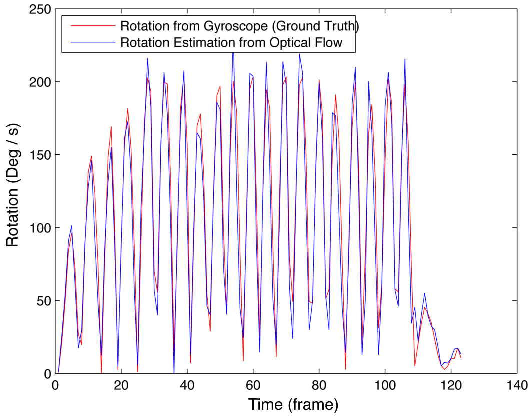
4.1.2. Shutter Angle and Rotation Estimation from a pair of Frames

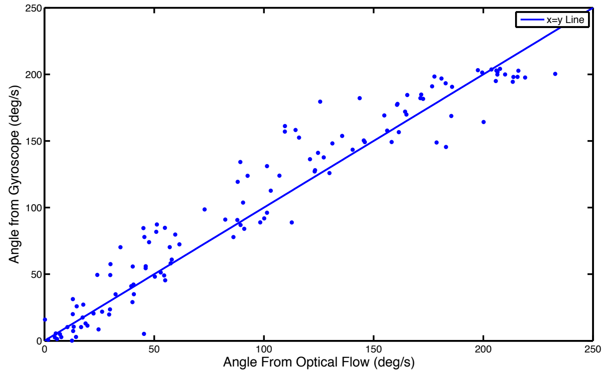
Figure 9 shows results from a synthetic sequence undergoing a series of varying rotations and with an animated shutter angle. Panel 1 in this figure shows the estimates for the magnitude of motion blur obtained from both the pair of frame method and single frame Klein and Drummond method, the latter being un-corrected for the known shutter exposure time. From this result it can be seen that in many cases where there is only a small amount of rotation, the single frame method from motion blur will over-estimate the amount of rotation that has occurred. However, the blur based system appears to consistently underestimate the value for rotation when there is a significant change in rotation, and this behaviour is to be expected - as detailed in Sec. 3.1, as the motion from blur will only represent a fraction of the frame time, whereas the frame to frame track will represent the full movement between frames.

Due to the noise in measuring rotation from blur, the resulting estimate for shutter angle is smoothed using a moving average filter (with a span of 4 frames) across the frame-set. This filtering is necessary because whilst the RANSAC algorithm described in Sec. 3.1 is able to reduce the effect of outlying estimates for rotation of the frame, certain conditions (further described in Sec. 5) will always produce incorrect results. The most significant source of error occurs when the magnitude of blur in the image is not sufficient for the accurate detection of the true change in focal length or rotation. By filtering these estimates we are able to reduce the impact of these errors whilst still maintaining an acceptable level of accuracy over periods where there is only a small amount of rotation present in the frame. A moving average filter was selected as this is a simple to implement filter that will filter out high-frequency changes in the estimate for shutter angle. We do not expect the shutter angle to change at every frame, so



Rotation with Gyroscope to Validate Rotation from Optical Flow Calculation (Poster)





Comparison of Rotation from Optical Flow Calculation with Gyroscope Data (Ground Truth) - Performed on the 'Poster' Real Dataset

This dataset was produced with a rigid camera-gyroscope rig in order to validate that the estimates produced by the optical flow algorithm for rotation in the presence of motion-blur were accurate when the rotation magnitude and axis of the camera is arbitrary and otherwise unknown.

Figure 8: Comparison of Results from Optical Flow based Rotation Estimation and Gyroscope Readings

this method allows for a single step change in shutter angle to be easily identified, whilst filtering the noisy calculation. Furthermore, outlying estimates that predict the shutter angle to be 1 or greater (i.e. the shutter was open longer than the frame time) are also automatically discarded.

4.2. Real Footage

The algorithms described in this work were tested over a set of real images captured by a Canon 700D SLR Camera along with a 70-200mm lens. The scenes shot were indoors and in good lighting conditions, and outdoors with natural light and some movement of objects in the scene (for example, trees moving in the wind and pedestrians walking through the frame). For the case of focal length estimation, a rotary encoder was attached to the lens barrel to track changes in rotation of the zoom wheel, and hence changes in the focal length. Each sequence consists of approximately 300 frames. In the case of rotation - the camera was rotated quickly and manually around an axis at various speeds and magnitudes, in order to produce a sequence that would exhibit large amounts of motion blur. Likewise, for changes in focal length, the zoom was also changed quickly and at varying speeds and magnitudes whilst filming. In all cases, the shutter speed was set to a constant 1/30th of a second - apart from the Chairs dataset where it was changed to 1/60th of a second after approximately 160 frames.

4.2.1. Shutter Angle and Rotation Estimation from a Pair of Frames

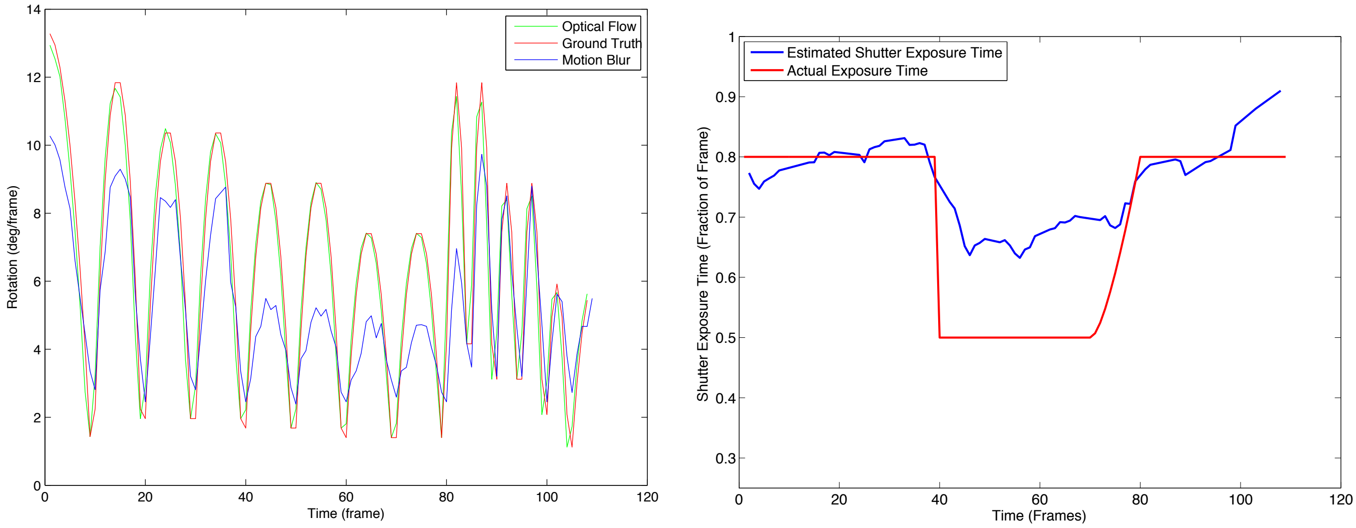
In order to validate the results produced using the 2 frame optical flow based method for determining camera rotation, the estimates obtained using this method on real footage were compared with the results obtained from a gyroscope rigidly attached to the camera during rotations around an axis. Figure 8 shows the results of this test. Ideally, the line plot for the angle estimated from optical flow against the gyroscope data should be identical, and the scatter plot for this data tend to an $x = y$ line.

Shown in figure 9 are the results obtained from rotating a camera around an axis over various magnitudes, and estimating rotation from both optical flow and blur. During shooting, the camera's shutter speed was changed from 1/30th of a second (0.83 of a frame at 25fps) to 1/60th of a second (0.415 of a frame at 25fps). Figure 9 also shows the estimated shutter angle as a fraction of the frame from the difference in estimations. As with the results from synthetic sequences, the value for shutter angle was calculated from a smoothed estimate for rotation from blur at each location above a threshold value.

4.2.2. Focal Length Change

Presented in figure 11 are the results for determining a change in focal length using a single frame using the method described previously. As with rotation from blur, the single frame method of determining focal length change

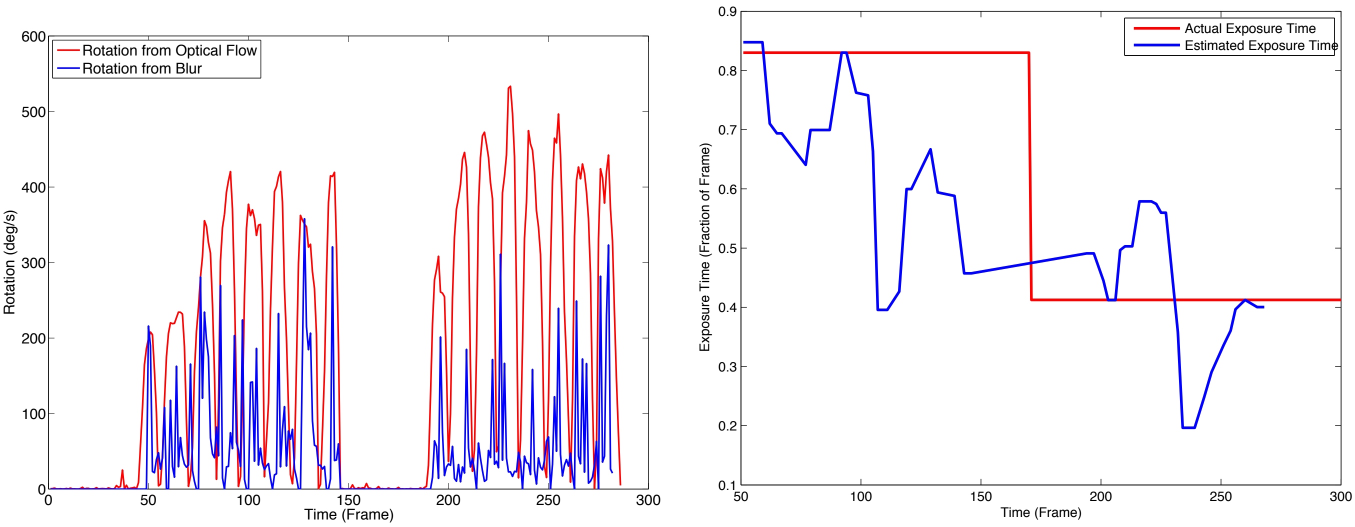
Synthetic Sequence (Synthetic Box)



(i) *(ii)*

Shutter Angle and Rotation Estimates from a Synthetic Dataset (Synthetic Box Sequence)

Rotation with Changing Shutter Angle. The final two frames above have a shutter angle of half the first three. (Chairs)



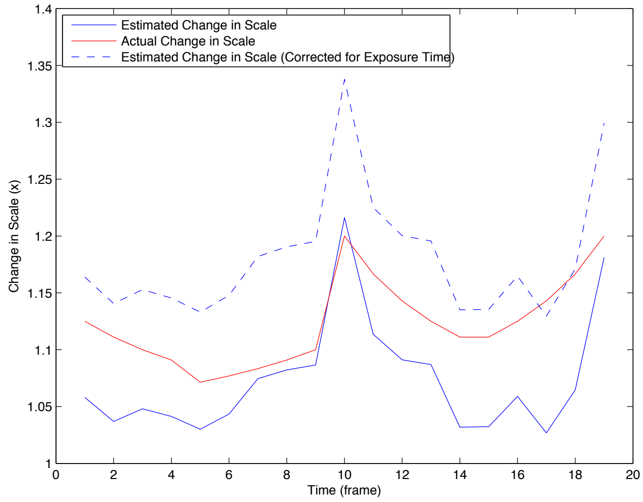
(iii) *(iv)*

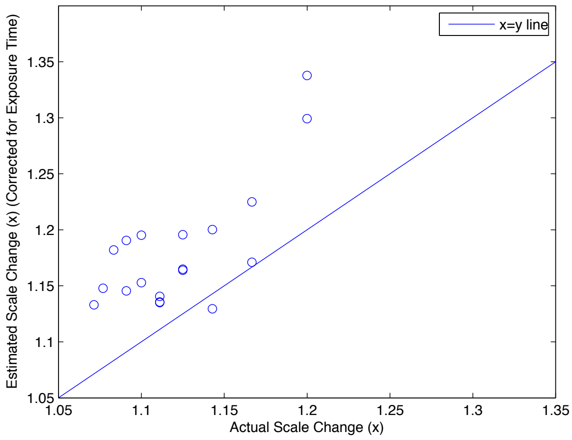
Shutter Angle and Rotation Estimates from a Real Dataset ('Chairs' Sequence)

Figure 9: Results for Estimating Rotation and Shutter Angle from Blur and Optical Flow

(a
b

Synthetic Focal Length Change (Zoom Synthetic Boxes)



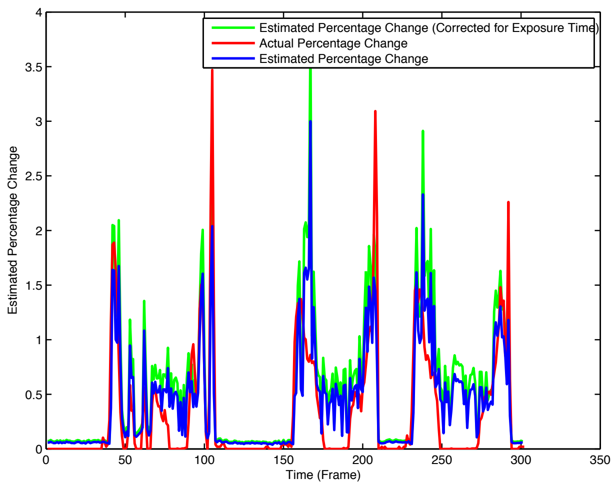


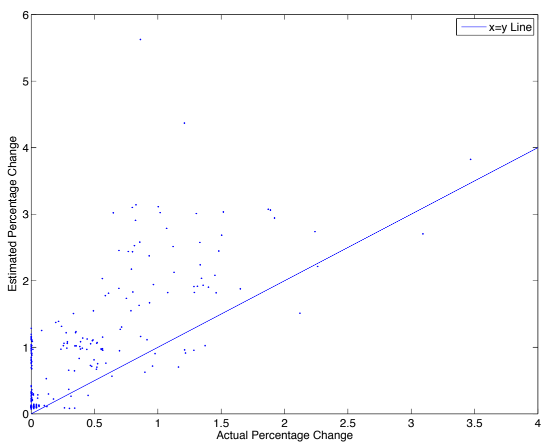
(i) *(ii)*

Change in Focal Length Estimates from a Synthetic Dataset. Ideally, the dashed-blue and solid-red lines in the left-hand chart should align, and the scatter plot should tend to an $x = y$ line.

Figure 10: Results for Estimating Change in Focal Length from Blur with a Synthetic Data Set

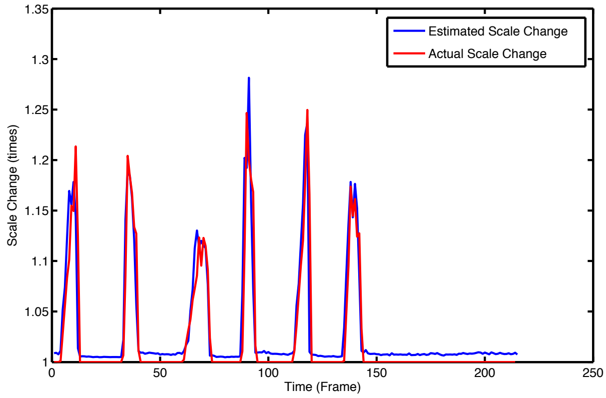
Real Focal Length Change (Zoom Boxes)



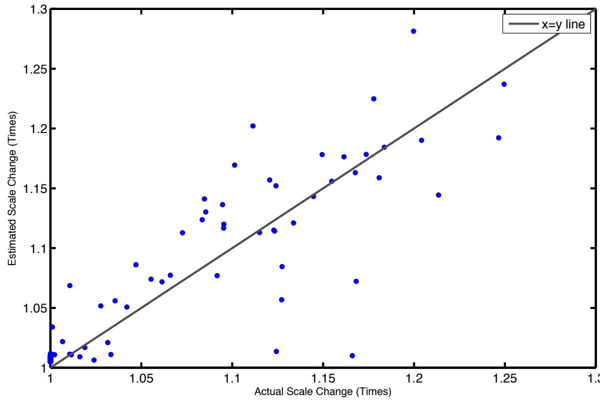


(i) *(ii)*

Change in Focal Length Estimates from a Real Dataset ('Zoom Boxes' Sequence). Ideally, the green and red lines in the left-hand chart should align, and the scatter plot should tend to an $x = y$ line.



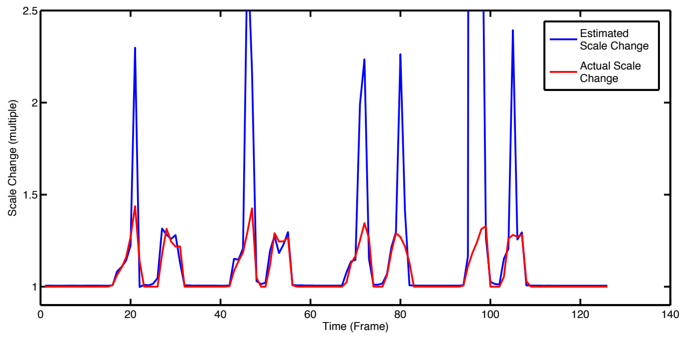
(iii)



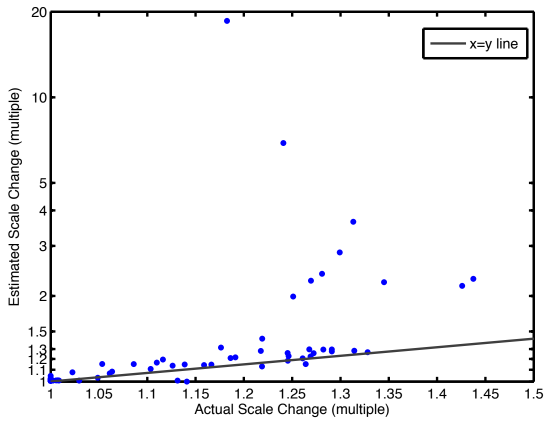
(iv)

Change in Focal Length Estimates from a Real Outdoor Dataset (‘Building’ Sequence). Ideally, the blue and red lines in the left-hand chart should align, and the scatter plot should tend to an $x = y$ line.

Figure 11: Results for Estimating Change in Focal Length from Blur for Real Datasets



(i)



(ii)

Change in Focal Length Estimates from a Real Outdoor Dataset with a Low Depth of Field Set (approx. 1.5m) (‘Flower’ Sequence). Ideally, the green and red lines in the left-hand chart should align, and the scatter plot should tend to an $x = y$ line.

Figure 12: Results for Estimating Change in Focal Length from Blur on a scene with a low Depth of Field

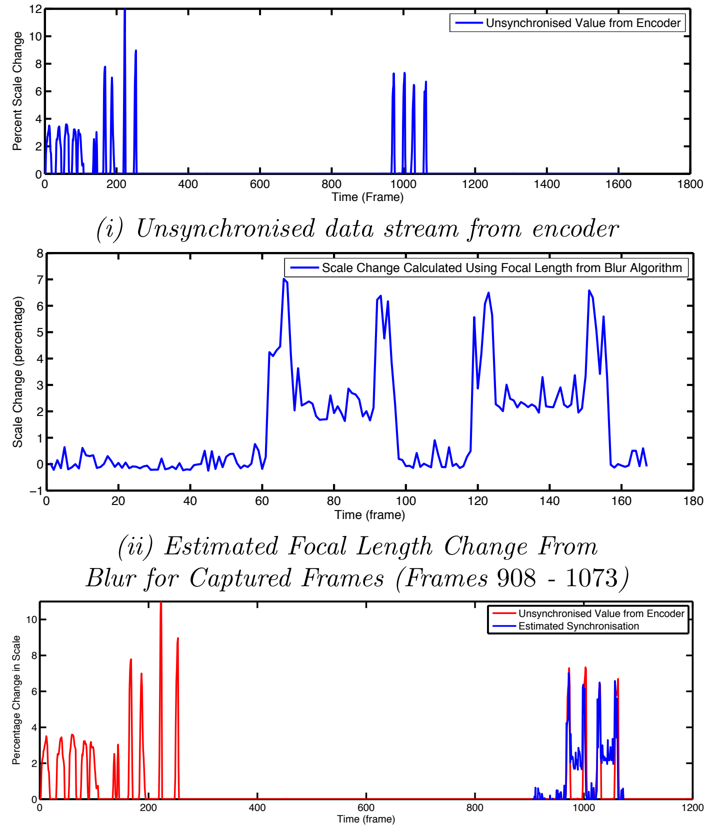
is unable to determine the direction of the change, hence data from the zoom encoder (taken as the ground truth) is converted to an absolute change in value. The initial indoor footage - 'Zoom Boxes' sequence in Fig. 11 panels *(i)* and *(ii)* was shot with good lighting conditions, however it can be seen that there is a smaller amount visual texture in the image, such as sharp edges and high contrast, when compared to the outdoor 'Building' sequence (panels *(iii)* and *(iv)* of the same figure). The result set for the 'Building' sequence as shown in panels *(iii)* and *(iv)* of Fig. 11 are clearly of a higher quality, and would suggest that the presence of good visual texture and a large number of sharp edges in the scene is important for achieving accurate results.

4.2.3. Alignment of Sensor Data with Video Footage

During capture of real data using both the gyroscope and zoom encoder equipment, it was necessary to synchronise the recording equipment with the video frames. This is performed by showing the camera a 'digislate' - a device which displays a time-code which refreshes at the specified framerate at the start of recording, and synchronising electronically this time-code with the data recording equipment. When the video is retrieved, the frames are manually inspected to read the time-code displayed on the device and correlate with the frame number of the sequence. Whilst this is a straightforward process to perform in a controlled environment, it is not practical in every shooting environment, e.g. if shooting from an aircraft. In such cases, manually aligning the data to the frame can be a difficult process. If an estimate can be found from frames with motion blur present as to the change in either zoom or rotation, then it could be used to assist in the alignment of the data in the case of failed synchronisation. One such way of achieving this would be the use of cross-correlation over both signals (estimate from blur and ground truth from sensors). Shown in figure 13 are the results from using the method of focal length estimation described in this work to align data from the zoom encoder sensor, compared to the actual synchronised values. In this case, the zoom encoder started recording positions before the camera started recording frames (recording changes in zoom that were not filmed) - shown in panel *(i)* of Fig. 13 and continued recording after the camera was stopped. The algorithm for estimating the amount of blur was run on the captured footage the results of which are shown in panel *(ii)* of the same figure and the data aligned using the results from the algorithm and cross correlation with the unsynchronised stream of data, the predicted alignment shown in panel *(iii)*. This predicted synchronisation shift differs by 1 frame from the actual known value of 908 frames.

4.3. Evaluating Algorithm Efficacy vs. Amount of Blur Present

Section 3.4 describes the method used for determining the amount of blur present in a scene, and shown here are



(i) Unsynchronised data stream from encoder

(ii) Estimated Focal Length Change From Blur for Captured Frames (Frames 908 *-* 1073*)*

(iii) Scale shift calculated from cross correlation of zoom estimates and unsynchronised stream. Shift estimated as 907 *frames*

Figure 13: Results for Estimating the Synchronisation between Camera and Zoom Encoder

the results for determining this metric (r_{blur}) along with the accuracy of the zoom estimates from Sec. 4.2.2. In order to evaluate the amount of blur necessary in an image to produce an accurate result, we calculate the amount of blur present in each frame of the sequence of real images using the method described in Sec. 3.4, where each frame has undergone a change in focal length of varying magnitude (including zero). This magnitude of blur is then compared to error between the estimate of scale change and the ground truth values for scale change at that frame. Figures 14,15 and 16 show the results of this analysis for each of the real datasets presented in Sec. 4. We would expect to see a higher proportion of over-estimates for the magnitude of scale change in the image, particularly at a low known scale change. The graphs for this analysis tend to support this conclusion - however, in all three cases there appears to be a reasonable amount of error when the scale change is greater than zero - but the amount of blur present in the image is not at it's maximum. In the graphs of figures 14,15 and 16, this can be seen as a reported under-estimate towards the middle of the blur-ratio scale (the x axis) where the red true scale-change line rises. This result would further support the conclusion that as a condition of a scale change being accurately estimated, it must cause significant motion blur in the image. However, it would appear that at the higher end of the scale change the method clearly over-estimates the true scale change by a considerable amount, and can sometimes under-report it. This would appear to contradict the theory that larger scale changes, resulting in larger amounts of blur present in the image (reflected by the rise of r_{blur}) should result in more accurate predictions using this method.

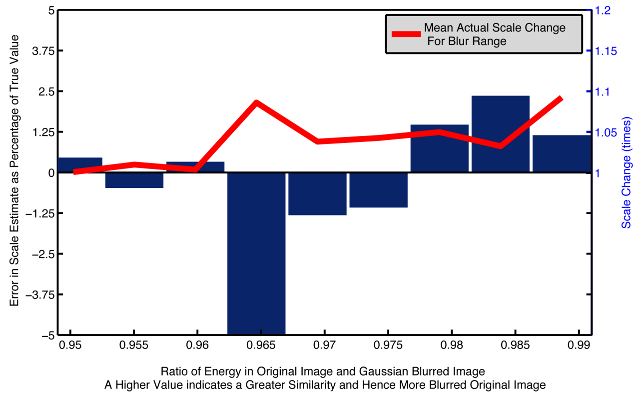


Figure 14: Results for comparing amount of blur in a frame with scale change estimate accuracy for the *'Boxes'* Dataset

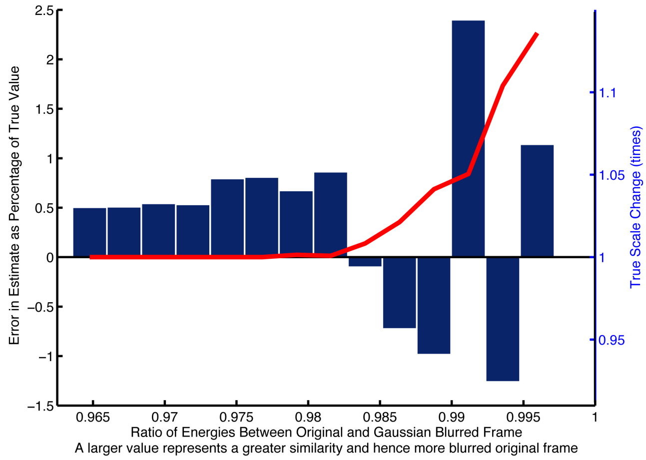


Figure 15: Results for comparing amount of blur in a frame with scale change estimate accuracy for the *'Building'* Dataset

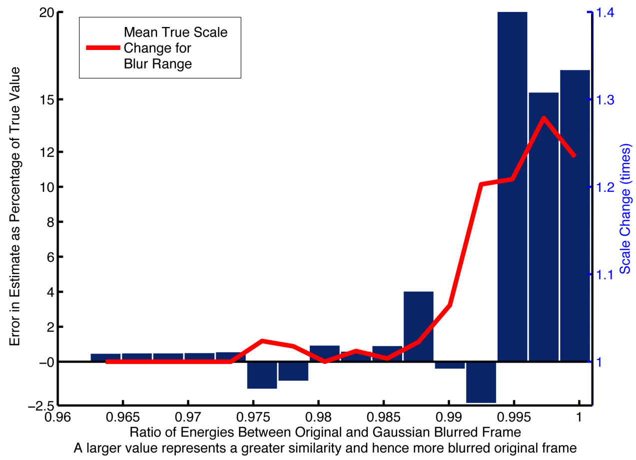


Figure 16: Results for comparing amount of blur in a frame with scale change estimate accuracy for the *'Flower'* Dataset

Using these results, it is proposed that a 'confidence' value of the estimated result can be predicted, in that for a images with a range of values calculated for r_{blur}, the expected result from using the original method for scale change from blur would be accurate to within a certain percentage error. This value could then be used to increase the accuracy of further results obtained from the same scene, in a situation where a ground truth would not be available. This would be especially useful in order to be able to categorise frames in which the scale change is likely to be zero, and hence saving the need to attempt to calculate a transform estimate for this frame. Applying the error metrics determined for the 'Building' scene to further footage of this scene (with the camera at a slightly different orientation) produces the results shown in Fig. 17 and Fig.18. These results are obtained by calculating the blur ratio (r_{blur}) from each frame and producing a 'corrected' result for this frame by applying the error metric for the range in which r_{blur} for this frame sits to the initial estimate. That is, if the frame is judged to have a value for r_{blur} as 0.987, the corrected result will be the estimated result scaled up by the error for this blur ratio from Fig. 15. If a value for r_{blur} is encountered that is not present within Fig. 15, then the value for scale change produced by the original algorithm is used. Similarly, if the value for r_{blur} is below a threshold indicating that no scale change is taking place, the corrected value is clamped to 0. We find that the cross correlation coefficient between the naive, raw estimates and the actual values to be 0.865, whereas the correlation coefficient between the corrected set and true values to be slightly better at 0.879.

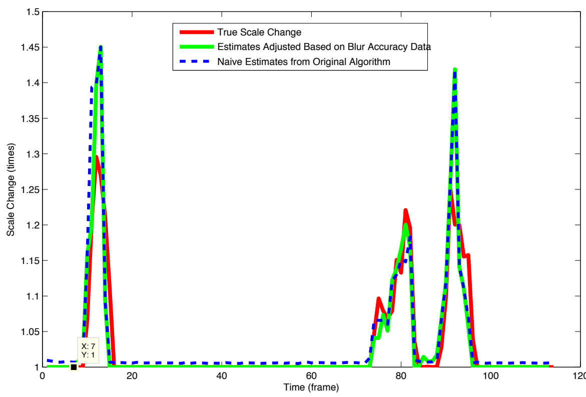


Figure 17: Comparison between the 'Naive' Focal Length from Blur Algorithm, and the 'Blur Aware' Method that multiplies results from the Naive Method with Error Factors Determined in Section 4.3. Ideally, the green line should be identical to the red, and closer to this than the blue line. Frames that are determined to have no scale change (a blur-ratio of less than 0.981) are capped at zero.

4.4. *Effects of Depth of Field*

Figure 12 shows the result of a real scene with a low depth of field (the *'Flower'* dataset). The focal distance in this scene was set to approximately 1.5m, whereas in the other real scenes used in this work, the focal distance is set to infinity. It can immediately be seen in panels *(i)* and *(ii)* that the results are somewhat more inaccurate than those from other images, with a tendency to greatly overestimate the true extent of scale change during large changes in scale. Images with a low depth-of field would typically have more blur in the frame regardless of motion blur introduced by scale change during shutter opening. This is something that Fig. 16 would confirm - as the zero, or close to zero scale change extends further along the blur ratio scale than in the results shown for other sequences. In theory, as long as part of the image is in focus, and this part has enough visual texture - such as sharp lines, then these would be blurred by the scale change and not from defocus - and could be used to calculate the scale change. In practice however, it is often the case that the in-focus part of the image would be at the centre of the image. As discussed in Sec.3.1, it is likely that points towards the centre of the image will be minimally scaled - and therefore unlikely to give a reliable estimate for the focal length change.

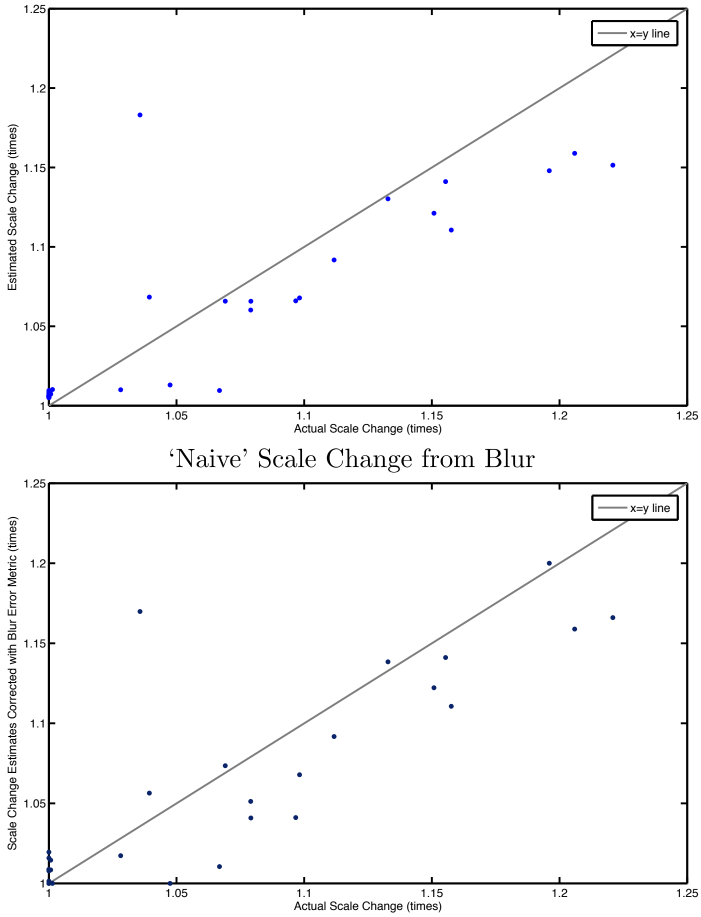


Figure 18: Comparison between the 'Naive' Focal Length from Blur Algorithm, and the 'Blur Aware' Method. Ideally, the points should tend to an x=y line, and the blur aware method should have points closer to this than the 'naive' method

5. Limitations

The results obtained from using motion blur in this work do suffer from several of the limitations discussed in the original Klein & Drummond paper. Notably, one of the most significant problems encountered for the estimation of parameters using blur is the need for a reasonable amount of blur to be present in order to be successfully detected. We have however presented a viable method to overcome this limitation somewhat by using prior knowledge of the error of the scale change estimate for a scene, and the amount of blur present in an image in order to better predict the scale change.

Another significant issue with the use of a single motion-blurred frame to estimate parameters is the inability of the system to cope with frames that have undergone more than one transformation - e.g. a rotation alongside a change in focal length. Another significant limitation of this work is the inability of the system to cope with large movement of objects in the scene. Our results suggest that a small amount of movement, such as pedestrians in a scene or a tree blowing in the wind will still allow for accurate results to be obtained. However, experimentation has shown that if the scene is completely obscured by movement, such as a vehicle passing in front of the camera during a focal length change, will cause the algorithm to fail.

Other limitations described in [6] for estimating parameters from blur are also present in this system, such as the intolerance to strobing, over-saturation, the requirement for pure rotation and a constant centre of rotation. However, when combined with the optical flow method described in [17], it is possible to determine the 'sign' of the rotation estimates. The method presented in [17], whilst extremely accurate (as shown by fig. 8), does have a significant limitation of requiring a large amount of resources to compute - often necessitating frames to be re-scaled prior to calculation. On average, for each blurred pair of frames at at size of 640×480 pixels, it would take approximately 30 seconds to compute an estimate for the optical flow, whereas the methods from blur would compute a result in near real-time on the same hardware (≈ 30 m/s), although this speed is highly dependent on the number of edgel sites selected and also the size of the image. Recent works in [2] and [3] have attempted to address this limitation.

Another factor that may have an effect on the result obtained for real footage would be the differences in blur introduced into a frame by a camera's rolling shutter (detailed in [8]). All of the algorithms described and used in this paper operate under the assumption that when a frame is blurred due to motion, the blur is always assumed to be constant across this frame. In a camera with a rolling shutter, each line of the sensor in the camera is sampled sequentially at different times. Therefore, during fast movement, in a camera with a rolling shutter, this assumption that all parts of the image will be blurred by a constant amount cannot be true. Investigating the impact and ways of minimising these effects in the algorithms using blur would be an important next stage of research.

6. Conclusions

This paper has shown an earlier method for determining changes in focal length during a single motion blurred frame, and has produced promising results from this method that allows for the estimates to be calculated quickly. We have also extended and combined two previous works in order to estimate the shutter angle of a frame. We have extended upon this work by presenting a new method to work with the original as part of an extended system in order to address previous limitations and enhance the accuracy of this new algorithm. We have also tested our methods on a new real data set and have been able to demonstrate that this improved method gives more accurate results, furthermore, we have examined how this system might cope with an image sequence with a shallow depth of field - and have uncovered potential limitations that this may present. An area of further research would be extending this system to handle frames which have been blurred by more than one type of motion - such as in the case of a translation and rotation, and work into this topic is ongoing.

[1] Alastair Barber, Matthew Brown, Paul Hogbin, and Darren Cosker. Estimating camera intrinsics from motion blur. In *Proceedings of the 11th European Conference on Visual Media Production*, CVMP '14, pages 6:1–6:10, New York, NY, USA, 2014. ACM.

[2] Xiaogang Chen, Jie Yang, Qiang Wu, Jiajia Zhao, and Xiangjian He. Directional high-pass filter for blurry image analysis. *Signal Processing: Image Communication*, 27(7):760 – 771, 2012.

[3] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. In *ACM SIGGRAPH Asia 2009 Papers*, SIGGRAPH Asia '09, pages 145:1–145:8, New York, NY, USA, 2009. ACM.

[4] Karen Goulekas. *The VES Handbook of Visual Effects*, chapter Postvis, pages 62–66. Elseveir, 2010.

[5] Alexandre Karpenko, David Jacobs, Jongmin Baek, and Marc Levoy. Digital video stabilization and rolling shutter correction using gyroscopes.

[6] Georg Klein and Tom Drummond. A single-frame visual gyroscope. In *Proc. British Machine Vision Conference (BMVC'05)*, volume 2, pages 529–538, Oxford, September 2005. BMVA.

[7] Wenbin Li, Yang Chen, JeeHang Lee, Gang Ren, and Darren Cosker. Robust optical flow estimation for continuous blurred scenes using rgb-motion imaging and directional filtering. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2014.

[8] Chia-Kai Liang, Li-Wen Chang, and H.H. Chen. Analysis and compensation of rolling shutter effect. *Image Processing, IEEE Transactions on*, 17(8):1323–1330, Aug 2008.

[9] Huei-Yung Lin. Vehicle speed detection and identification from a single motion blurred image. In *Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1. Seventh IEEE Workshops on*, volume 1, pages 461–467, Jan 2005.

[10] S.K. Nayar and M. Ben-Ezra. Motion-based motion deblurring. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):689–698, June 2004.

[11] T. Okatani and K. Deguchi. Robust estimation of camera translation between two images using a camera with a 3d orientation sensor. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 275 – 278 vol.1, 2002.

[12] Oliver J. Woodman. An introduction to inertial navigation. Technical Report UCAM-CL-TR-696, University of Cambridge, August 2007.

[13] Richard J. Radke. *Computer Vision for Visual Effects*. Cambridge University Press, 2013.

[14] Ioannis M. Rekleitis. Steerable filters and cepstral analysis for optical flow calculation from a single blurred image. In *In Vision Interface*, pages 159–166, 1996.

[15] Jianbo Shi and Carlo Tomasi. Good features to track. In *1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593 – 600, 1994.

[16] Yu-Wing Tai, Hao Du, M.S. Brown, and S. Lin. Image/video deblurring using a hybrid camera. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.

[17] Li Zhang, T. Portz, and Hongrui Jiang. Optical flow in the presence of spatially-varying motion blur. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 0:1752–1759, 2012.