

© 2011 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Vision Based Localization under Dynamic Illumination

Jared Le Cras, Jonathan Paxman, Brad Saracik

Department of Mechanical Engineering
Curtin University
Perth, Australia

Abstract—Localization in dynamically illuminated environments is often difficult due to static objects casting dynamic shadows. Feature extraction algorithms may detect both the objects and their shadows, producing conflict in localization algorithms. This work examines a colour model that separates brightness from chromaticity and applies it to eliminate features caused by dynamic illumination.

The colour model is applied in two novel ways. Firstly, the chromaticity distortion of a single feature is used to determine if the feature is the result of illumination alone i.e. a shadow. Secondly, the chromaticity distortion of features matched between images is examined to determine if the monochrome based algorithm has matched them correctly.

These two applications are put through a variety of tests in simulated then real world environments to assess their effectiveness in dynamically illuminated scenarios. The results demonstrate a significant reduction in the number of feature mismatches between images with dynamic light sources. The evaluation of the techniques individually in a Simultaneous Localization and Mapping (SLAM) task show substantial improvements in accuracy, with the combination of the two techniques producing a localization result that is highly robust to the environmental lighting.

Keywords-localization; mining; dynamic illumination;

I. INTRODUCTION

Dynamically illuminated environments pose a challenge to vision based localization and mapping that is often overlooked due to the prevalence of static light sources in environments used for evaluation. Scenarios exist, however, where static light cannot be guaranteed. These include applications such as slow moving equipment affected by the movement of the sun, aerial photographs taken at different times of day and the area of focus for this work: underground mining environments. We found that algorithms, such as Scale Invariant Feature Transform (SIFT) and Features from Accelerated Segment Test (FAST) corner detection, frequently extract shadow boundaries as potential features for tracking. This is not a problem within statically illuminated environments, where shadow features are as reliable as the objects that cast them. However, when the primary light source is dynamic, the inconsistency of stationary objects and moving shadows causes difficulties for localization algorithms.

This work examines an approach to distinguish object based features from shadow based features in order to improve localization by tracking only those considered to be robust to variations in illumination. It also examines a technique to reduce mismatches produced by monochrome based matching algorithms in dynamically lit scenes by utilizing colour information. This work is motivated by our efforts in the surveying and mapping of underground mine systems [12, 10].

II. RELATED WORK

Vision based localization has been an area of active research for the last 20 years. Successful localization has been realized by many research groups [17, 26, 25, 15, 6]. However, all of the scenarios used to evaluate these localization algorithms have involved a dominant static light source and therefore the impressive results cannot be transferred to scenarios that involve a primary dynamic light source. To improve the robustness of vision based localization to dynamic illumination, the properties of the features used for localization must be examined. Two feature extraction techniques that are prevalent in vision based localization algorithms are Scale Invariant Feature Transform (SIFT) by Lowe [13] and Features from Accelerated Segment Test (FAST) corner detection by Rosten and Drummond [16]. SIFT has been proven to be the most robust technique for extracting and matching features across multiple images [14], whereas FAST corner detection is designed to work with real-time video frame rates. Filtering the features extracted by these two techniques can improve localization as well as partially improving illumination robustness.

SIFT is scale invariant and partially invariant to orientation and slight changes in illumination [13]. The most prevalent problem experienced by those that use SIFT for feature extraction and localization is the large number of features extracted from each image. Improving the efficiency of feature matching between images is still an area of active research. Brown *et al.* [3] group ‘interest points’ together to form complex features. Matching of these features between frames produces a transformation prediction that allows the identification and removal of outlier matches. The approach improves orientation invariance; however, only limited improvement in illumination invariance is reported. The technique applied by Tamimi *et al.* [24] is to search only for a specific ‘blob’ feature type that matches well between images. These blob features are areas of the image that are significantly

lighter or darker than their surroundings. Any feature detected that is determined to not fall into the blob category is discarded. A limit on the maximum number of extracted features is also imposed to restrict matching time. Again, only limited improvement to illumination robustness is reported.

FAST corner detection can produce a smaller number of features than SIFT and is consequently faster and more suitable for real time implementation. A region is determined to be a feature if a centroid pixel is significantly darker or lighter than the surrounding circumference of 16 pixels. The low number of features has led to several attempts to improve matching between images. Klein *et al.* [11] improve matching robustness to image blur by using features at multiple pyramid levels. However, this pyramidal approach produces no improvement in illumination robustness. Chekhlov *et al.* [5] improve FAST corner detection matching through the use of camera pose estimation. A prediction of feature locations based on a movement model is used to reduce search regions for matching and improve matching robustness by reducing the weighting of dynamic illumination based features. This technique is applied by Civera *et al.* [6] in a (six degree of freedom) monocular EKF SLAM algorithm. The filtering techniques developed in our work are evaluated with a localization algorithm based on Civera's.

The limiting factor of these modest improvements to illumination robustness is the use of monochrome images. Colour information has been successfully exploited in other areas of computer vision research to improve illumination invariance. Swain *et al.* [23] were the first to use a three dimensional RGB histogram to describe pixel colour values for recognition, known as 'colour indexing'. Finlayson *et al.* expanded upon this approach to produce basic, lighting invariant, object recognition using 'colour angles' [7]. They demonstrated improved efficiency over colour indexing, at the cost of number of correct recognitions. An alternative approach trialed by Geusebroek *et al.* estimated the original colour of an object mathematically given known lighting conditions [8]. This technique was robust to viewing direction, surface orientation, highlights, illumination direction, illumination intensity, illumination colour and inter-reflection, but was limited to a range of known materials and strictly controlled lighting conditions. Successful shadow detection in a dynamic environment was achieved by Horprasert *et al.* [9] to detect moving objects against a static background. An RGB histogram compared successive pixel colours between frames to produce a colour model based on two components: brightness and chromaticity. These two components were used to examine a change in pixel colour over time such that a differentiation could be made between static background, shadow and dynamic foreground. The concept behind the technique is that a shadow cast on an object will result in a large change in brightness, whereas the chromaticity difference between pixels will remain small. Our approach to the filtering of SLAM features is based on this brightness and chromaticity model and will be examined in detail in Section III.

The use of colour information to improve feature extraction algorithms is not a novel concept. Both SIFT and FAST corner detection use only the monochrome version of an image, so it has long been assumed that the wealth of information available

in the colour image could be used to improve descriptors and matching. Abdel-Hakim *et al.* [1] combine both colour and geometrical information in the feature description, resulting in an increased number of extracted and matched features. Burghouts *et al.* [4] integrate the Gaussian opponent colour model into SIFT; this model consists of intensity, red-green and yellow-blue channels. The separation of intensity enhances illumination invariance and results in improved matching when compared to standard SIFT over a range of lighting conditions. Unfortunately, there is no efficiency evaluation since the algorithm was designed for object recognition, not real-time localization.

The colour model employed by this paper in Section III also separates intensity (brightness) from the colour channels. We derive two novel techniques which aid in the accurate use of extracted features for localization in dynamically illuminated environments.

There are alternative approaches to improving illumination robustness in localization as demonstrated by the promising results achieved by Silveira *et al.* [18, 19] by modeling illumination change as a surface that can evolve over time. The model is combined with projective geometry to produce basic planar visual tracking. This approach requires no a priori knowledge about the light source(s) or the subject surface materials. A planar feature in an image is 'warped' to match the same feature seen from a different viewpoint in a subsequent image. The 'warp' required to match the features is then used to perform basic localization. The need for high quality planar features is the limiting factor in this technique, despite its illumination robustness, so therefore would not be applicable to an underground mining environment where there are limited planar surfaces.

Alternative approaches to overcoming errors due to illumination changes include combining visual information with depth information (SungHwan *et al.* [21, 22]) or providing the localization algorithm with a priori information (Bischof *et al.* [2], Steinbauer [20]). Our work avoids the use of depth information to remove the need for additional sensors and sensor fusion. The need for a priori information was also avoided as it hinders real world application.

The lack of a robust localization technique for dynamically illuminated environments led to the development of the technique detailed in the following section. This technique uses the colour information associated with extracted image features to improve robustness to lighting while requiring only a stream of colour images from a single camera.

III. TECHNIQUE DEVELOPMENT

In order to correctly distinguish between physical features and illumination artifacts, it is desirable to know the true colour of objects (as opposed to the perceived colour). The difficulty in determining the true colour of an object is due to the large number of factors that affect perceived colour [8]. An approximation of true colour is provided by the brightness and chromaticity model proposed by Horprasert *et al.* [9]. The model compares the distortion between two pixel colours in RGB space. A line OE_i passing through the origin and the first pixel colour ($E_i = [E_R(i), E_G(i), E_B(i)]$) is called the expected

chromaticity line and is used to determine the distortion in chromaticity and brightness. If the second pixel colour ($I_i = [I_R(i), I_G(i), I_B(i)]$) is on this line, then there is a distortion in brightness only (α_i), otherwise there is also a distortion in chromaticity (CD_i), (see Fig. 1).

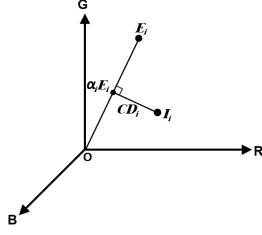


Figure 1. Brightness and chromaticity distortion between two colours (E_i, I_i).

The brightness distortion (α_i) is a scalar value that represents the point on the expected chromaticity line that is closest to the comparative colour (I_i). It is obtained by minimizing:

$$\varphi(\alpha_i) = (I_i - \alpha_i E_i)^2 \quad (1)$$

Brightness distortion will be exactly 1 if the comparative colour has the same brightness, less than 1 if it is darker and greater than 1 if it is brighter.

Chromaticity distortion (CD_i) is the orthogonal distance between the expected chromaticity line and the comparative colour (I_i). The chromaticity distortion of pixel i is given by:

$$CD_i = \| I_i - \alpha_i E_i \| \quad (2)$$

Our approach applies this colour model to features extracted from images rather than individual pixels. Both SIFT and FAST corner detection are used as feature extraction algorithms. SIFT is used for offline applications due to its popularity in machine vision and robustness [14], whereas FAST corner detection is used for the real time localization due to its speed. A 3x3 pixel region is used to represent the colour information of features extracted by either technique. SIFT has scale invariant features, so the 3x3 pixel region is scaled to the magnitude of the extracted feature. In order to maintain the 3x3 pixel colour information, the feature is divided into a 3x3 grid and the mean colour of the pixels in each grid square are stored as a single pixel representative. FAST corner detection uses a 3 pixel radius window having a 16 pixel circumference. A 3x3 pixel region is again used to characterize the colour of the center of the extracted feature. Since FAST features are not scaled, this simple approach works for all extracted features. The colour of a feature (E_i) is represented by the mean RGB values of the 3x3 pixel region i.e.

$$E_i = [\mu R(i), \mu G(i), \mu B(i)] \quad (3)$$

The standard deviation of the RGB values for each feature is given by:

$$s_i = [\sigma R(i), \sigma G(i), \sigma B(i)]. \quad (4)$$

The standard deviation is then used to calculate the signal-to-noise ratio. The signal-to-noise ratio is applied differently by each of the two techniques derived in this work. These

techniques will be detailed individually in the following two sections.

A. Removal of Shadow features

The first novel technique applied in this work is the use of the colour model from the previous section to detect features extracted from the edges of shadows. Dynamically illuminated environments contain dynamic shadows. The tracking of dynamic shadows will lead to erroneous localization. Therefore, to improve illumination robustness, features extracted from these shadows must be identified and removed. Both SIFT and FAST corner detection extract features from the ‘edges’ of objects in the image. These edges are therefore the focus of identifying a feature as shadow-based or object-based.

To determine the likelihood of a feature being extracted from a shadow, the chromaticity distortion of the feature is calculated. The RGB value of each pixel in the 3x3 pixel feature window is compared to the mean colour of the feature. The resulting range of chromaticity distortions is then compared to a threshold to identify a feature as shadow-based or object-based. A feature extracted from the edge of a shadow will have a low chromaticity distortion due to the ‘edge’ containing a variation in illumination intensity only. A feature extracted from the edge of an object will usually have a higher chromaticity distortion due to the edge containing a variation in colour.

Brightness and chromaticity distortion is determined by comparing each of the 9 pixels in the 3x3 feature window (p_i) to the colour value of the overall feature.

$$\alpha_i = \frac{p_{Ri} \mu_R + p_{Gi} \mu_G + p_{Bi} \mu_B}{\left(\frac{\mu_R}{\sigma_R}\right)^2 + \left(\frac{\mu_G}{\sigma_G}\right)^2 + \left(\frac{\mu_B}{\sigma_B}\right)^2} \quad (5)$$

$$CD_i = \sqrt{\left(\frac{p_{Ri} - \alpha_i \mu_R}{\sigma_R}\right)^2 + \left(\frac{p_{Gi} - \alpha_i \mu_G}{\sigma_G}\right)^2 + \left(\frac{p_{Bi} - \alpha_i \mu_B}{\sigma_B}\right)^2} \quad (6)$$

Features that have a chromaticity distortion range below a selected threshold (τ_{SFR}) are removed as they are not considered to be robust to changes in illumination. The remaining features can then be used for matching, increasing the probability of illumination robustness during localization.

B. Colour Based Feature Matching

The second novel technique applied in this work is the use of the same colour model to increase the constraints on feature matching. The technique filters standard feature matching results and so can be applied to any existing feature matching algorithm.

A correctly matched feature will exhibit a small difference in true colour, as opposed to an incorrectly matched feature which will have a significant difference in true colour. Both SIFT and FAST corner detection use monochrome images for feature matching and therefore can be misled by lighting conditions that cause different objects to look similar in grey scale.

To identify a match as being of similar true colour, the chromaticity distortion between the matched features is calculated and compared to a threshold. The distortion value will be low, regardless of lighting conditions, if a correct match has been made.

A standard feature matching algorithm is used to produce a set of matched features. The brightness and chromaticity distortion is determined by comparing the normalized colour of one feature to the normalized colour of the matched feature.

$$\alpha_i = \frac{\frac{\mu_{1R}\mu_{2R}}{\sigma_{1R}\sigma_{2R}} + \frac{\mu_{1G}\mu_{2G}}{\sigma_{1G}\sigma_{2G}} + \frac{\mu_{1B}\mu_{2B}}{\sigma_{1B}\sigma_{2B}}}{\left(\frac{\mu_{1R}}{\sigma_{1R}}\right)^2 + \left(\frac{\mu_{1G}}{\sigma_{1G}}\right)^2 + \left(\frac{\mu_{1B}}{\sigma_{1B}}\right)^2} \quad (7)$$

$$CD_i = \sqrt{\frac{(\mu_{1R} - \alpha\mu_{2R})^2}{\sigma_{1R}\sigma_{2R}} + \frac{(\mu_{1G} - \alpha\mu_{2G})^2}{\sigma_{1G}\sigma_{2G}} + \frac{(\mu_{1B} - \alpha\mu_{2B})^2}{\sigma_{1B}\sigma_{2B}}} \quad (8)$$

Matched features that have a chromaticity distortion above a selected threshold (τ_{CBM}) are discarded as they are considered a false match due to lighting conditions.

IV. EXPERIMENTAL RESULTS

A. Comparison of Colour Models

The chromaticity distortion model has never before been applied to a localization task. To confirm that it was an appropriate choice of colour model, it was compared to two other models that separate colour from intensity – HSV and colour angles. Hue Saturation Value (HSV) represents colour using an angle to specify the base colour on a circular colour chart, then separately defines a saturation and ‘darkness’ value. Colour angles describe the difference between two colours in RGB space as the angle between the two RGB vectors. The image of a handle casting a shadow on a door (Fig. 2) was used to compare these three techniques by assessing their ability to differentiate between a true colour difference and a perceived colour difference. SIFT is used to identify about 3000 features in the image, each colour model is then tuned to reject around 1000 features based on a comparison of each pixel in the 3x3 grid to the mean feature colour, as described in Section III. The results in Figs. 3-5 show that the chromaticity distortion model is the only colour model to correctly reject the features on the edge of the shadow and on the highlight. HSV and colour angle produced similar results, both poorly identifying the shadow and highlight. The chromaticity distortion technique was trialed



Figure 2. Image used for colour model comparison test

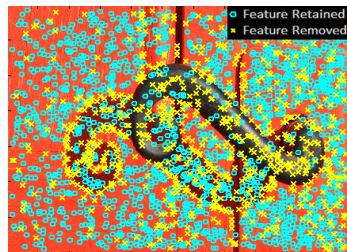


Figure 3. Features retained and removed by chromaticity model

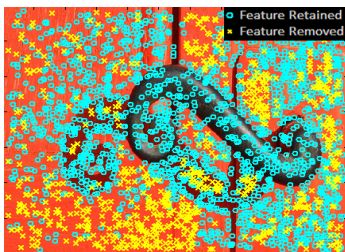


Figure 4. Features retained and removed by HSV model

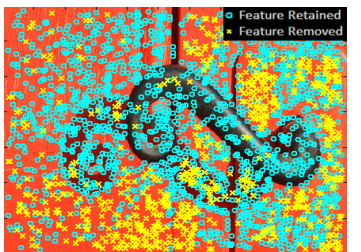


Figure 5. Features retained and removed by colour angles model.

on other images to test the repeatability of the results and similar outcomes were achieved.

B. Dynamic Illumination Localization (Simulated Scene)

To prove the concept further, a dynamic illumination scenario was introduced using computer generated images (created using POVray). This allowed explicit control over the camera and lighting conditions. A series of images depicting movement through a scene with dynamic lighting allowed full assessment of the effect on localization caused by chromaticity distortion filtering. The camera moved with a constant velocity through a scene containing a chess board. The motion path was a straight line parallel to the edge of the chess board, while the single light source moved from slightly behind the chess board to directly overhead, causing dynamic shadows (see Figs. 6 & 7).

The SLAM implementation for this experiment was based on the six degree of freedom monocular SLAM algorithm written by Civera *et al.* [6]. The algorithm was modified by the addition of the chromaticity distortion filter developed in this paper. Figs. 9-12 are a top view representation of the 3D localization produced by the monocular SLAM algorithm. Each image shows the localization path as a line starting at the origin (0,0) and finishing in the center of the triangle representing the final orientation of the camera. The ellipsoids in each image center on a tracked feature and represent the locational uncertainty of that feature. The remaining lines point to features that were observed in only a single image and hence have no depth estimate. The threshold used in each test was a constant and based on effective values used in previous experiments. All compensation techniques were assessed, starting with uncompensated (Fig. 9), shadow features removed (Fig. 10), colour based matching (Fig. 11) and finally both techniques combined (Fig. 12).

The ground truth for the test is shown in Fig. 8. The dimensions are unitless as the scene is computer generated. The path is a straight line with no change in camera orientation. The first test introduced the dynamic light source and was run using the unfiltered SLAM algorithm. The dynamic shadows have a significant influence on the effectiveness of the localization, resulting in the heavily distorted path seen in Fig. 9.

The second test applied the feature filtering approach outlined in Section III-A. This approach produced a significant improvement in localization, yet still contained evidence of the distortion caused by the dynamic light source (Fig. 10). The third test filtered feature matches based on their chromaticity

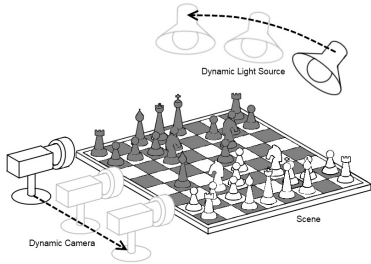


Figure 6. Dynamic illumination scene



Figure 7. Dynamic illumination scene

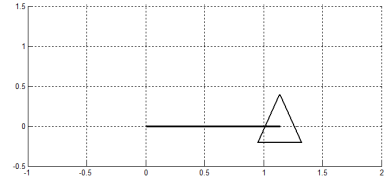


Figure 8. Ground truth

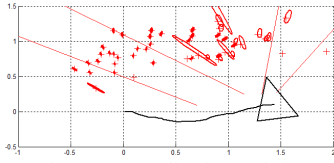


Figure 9. Localization path for dynamic illumination, SLAM algorithm only.

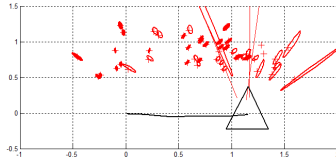


Figure 10. Localization path for dynamic illumination, with shadow features removed (SFR).

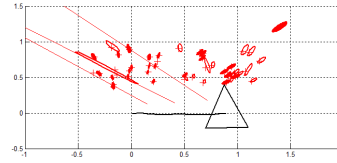


Figure 11. Localization path for dynamic illumination, with colour based matching (CBM).

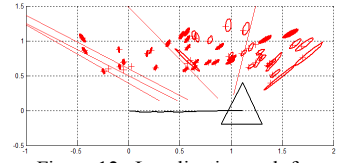


Figure 12. Localization path for dynamic illumination, with both techniques implemented (SFR & CBM).

distortion as detailed in Section III-B. The localization path produced by this approach completely removed the distortion caused by the dynamic light source. The consequence of this improvement was a large reduction in the number of maintained matches, resulting in the discrimination between the displacement of this path and the ground truth (Fig. 11).

The final test combined both filtering techniques. The resulting localization path (Fig. 12) demonstrated the effectiveness of the approach. The distortion produced by the dynamic light source was completely removed and the high level of maintained features produced a highly accurate displacement. These results demonstrate the effectiveness of the two novel techniques derived in this work as approaches to overcoming the challenges of localization in dynamically illuminated environments.

C. Dynamic Illumination Localization (Real World) with Automatic Thresholding

Computer generated scenes are a useful tool to rapidly prototype algorithms for filtering based on chromaticity distortion; however, it is difficult to completely simulate the imperfections of real image data, including uneven lens distortion, unfocused features, lens flare, graininess and poor contrast. We therefore continue by verifying the developed techniques in a real world scenario.

To analyze the real world performance of the algorithm, a dataset was recorded of movement through a scene with dynamic lighting (see Fig. 13). The camera moves through the scene sideways, five meters to the left in a straight line, the



Figure 13. An image from the real world with dynamic lighting dataset

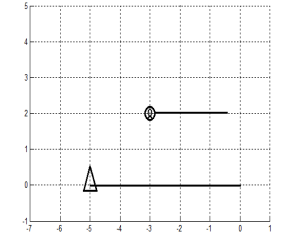


Figure 14. The path travelled by the camera (triangle) and light source (circle)

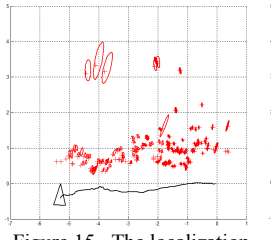


Figure 15. The localization path of the uncompensated algorithm.

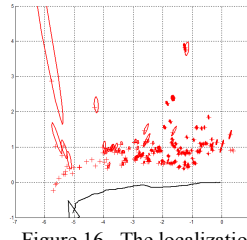


Figure 16. The localization path of the compensated algorithm, fixed threshold.

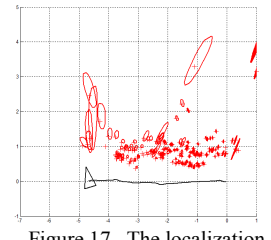


Figure 17. The localization path of the compensated algorithm, auto scaled threshold.

light source moves parallel to the camera yet at a slower speed producing shadows that drift away from the direction of camera movement (see Fig. 14). This dataset was run through the uncompensated monocular SLAM algorithm (Fig. 15), then through the compensated algorithm with a fixed threshold (Fig. 16) and finally through the compensated algorithm with an automatically scaled threshold (Fig. 17).

The localization results are again presented in the same format as previous results. Running the dataset through the uncompensated algorithm produced accurate localization for a brief window of time, as expected by the slight illumination invariance characteristic of most feature extraction techniques. As the dataset continues, however, the tracking of dynamic shadow features leads to the poorly localized path seen in Fig. 15. The results produced by the algorithm incorporating chromaticity distortion with a fixed threshold are also poorly localized. The use of a single fixed threshold for the entire dataset was too limiting to produce accurate results.

Images that are considered to be low in colour (or 'bland') need to have the SFR threshold relaxed otherwise all features will be classified as shadows. However, the CBM threshold must be stringent otherwise the colour difference between mismatched features will always pass through the chromaticity distortion filter. To overcome this discrimination, automatically scaling thresholds were implemented based on the colour content of an image. The chromaticity distortion colour model was again utilized to determine the colour range of an image. The chromaticity distortion of a sample of random pixels was compared to the mean colour of the entire image. The standard

deviation of the sample of random pixels correlated to the colour range present in each image and was used to scale the threshold.

Far superior results are produced by the algorithm incorporating chromaticity distortion based filtering with an automatically scaling threshold (Fig. 17). The localization path is highly correlated to the ground truth and the final camera orientation rotation is minimal.

These results prove the real world performance of chromaticity distortion based filtering as an approach to improving illumination robustness. The modularity of the approach allows it to be easily incorporated into feature extraction algorithms that are used in many real world applications.

V. CONCLUSION

This work has demonstrated substantial improvements in the robustness of localization under dynamic illumination by applying a chromaticity distortion model. These findings were the result of the successful integration of a modified version of the chromaticity distortion model with a fully implemented, six degree of freedom monocular SLAM algorithm. The two techniques resulting from this pairing were the removal of shadow based features and the colour based filtering of feature matching.

The application of these techniques has resulted in localization under dynamic illumination which is almost as accurate as localization under static illumination in simulation. Testing in a real world scenario shows significant reductions in localization distortion under dynamic illumination, resulting in a localization path which is close to ground truth.

In future work these techniques will be integrated into a hybrid SLAM system to improve localization and mapping for deployment in underground mining environments. As we have noted, there are a number of application areas involving dynamic illumination. These may include slow moving agricultural equipment, where the movement of the sun becomes a factor in localization, or feature tracking in aerial photographs where significant amounts of time may elapse between images, causing a dynamic lighting effect.

REFERENCES

[1] Abdel-Hakim, A., Farag, A. 2006. CSIFT - A SIFT descriptor with color invariant characteristics, IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 1978-1983.

[2] Bischof, H., Wildenauer, H., Leonardis, A. 2004. Illumination insensitive recognition using eigenspaces. *Computer Vision and Image Understanding*. Vol. 95, no. 1, pp. 86-104.

[3] Brown, M., Lowe, D. 2002. Invariant Features from Interest Point Groups. *British Machine Vision Conference*.

[4] Burghouts, G., Geusebroek, J. 2009. Performance evaluation of local colour invariants. *Computer Vision and Image Understanding* 113 pp. 48-62

[5] Chekhlov, D., Pupilli, M., Mayol, W., Calway, A. 2007. Robust real-time visual slam using scale prediction and exemplar based feature description. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*.

[6] Civera, J., Grasa, O., Davison, A., Montiel, J. 2010. 1-Point RANSAC for EKF Filtering: Application to Real-Time Structure from Motion and Visual Odometry. *Journal of Field Robotics*. Vol. 27 no. 5, pp. 609-631.

[7] Finlayson, G., Chatterjee, S., Funt, B. 1996. Color Angular Indexing. *Proceedings of the 4th European Conference on Computer Vision - Volume II - Volume II table of contents*, Pages: 16 - 27.

[8] Geusebroek, J., van den Boomgaard, R., Smeulders, A., Geerts, H. 2001. Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 23, no. 12, pp. 1338-1350.

[9] Horprasert, T., Harwood, D., Davis, L. 2000. A robust background subtraction and shadow detection. In *Proceedings of the Asian Conference on Computer Vision*.

[10] Jarosz, A., 2008. Development of Inspection System for Evaluation of Ore-passes at Grasberg Mine, PT Freeport, Indonesia. In *Proceedings, The 21st World Mining Congress & Expo, September 2008*.

[11] Klein, G., Murray, D. 2007. Parallel tracking and mapping for small AR workspaces. *Proceedings of the International Symposium on Mixed and Augmented Reality*.

[12] Le Cras, J., Paxman, J., Saracik, B., Jarosz, A. 2009. An inspection and surveying system for vertical shafts. In *Proceedings Australasian Conference on Robotics and Automation 2009*.

[13] Lowe, D. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* Vol. 60, no. 2, pp. 91-110

[14] Mikolajczyk, K., Schmid, C. 2005. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 27, no. 10, pp. 1615-1630.

[15] Mouragnon, E., Lhuillier, M., Dhome, M., Dekeyser, F., Sayd, P. 2009. Generic and real-time structure from motion using local bundle adjustment. *Image and Vision Computing*. Vol. 27, no. 8, pp. 1178-1193.

[16] Rosten, E., Drummond, T. 2006. Machine learning for high-speed corner detection. *Lecture Notes in Computer Science, 2006, Volume 3951/2006*, pp. 430-443.

[17] Se, S., Lowe, D., Little, J. 2001. Vision-based mobile robot localization and mapping using scale-invariant features. In *Proceedings IEEE International Conference on Robotics and Automation*. Vol. 2, pp. 2051-2058.

[18] Silveira, G., Malis, E. 2007. Real time visual tracking under arbitrary illumination changes. In *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1-6.

[19] Silveira, G., Malis, E., Rives, P. 2008. An efficient direct approach to visual slam. *IEEE Transactions on Robotics*. Vol. 24, no. 5, pp. 969-979.

[20] Steinbauer, G., Bischof, H. 2005. Illumination Insensitive Robot Self-Localization Using Panoramic Eigenspaces. *Lecture Notes in Computer Science*. Vol. 3276, pp. 84-96.

[21] Sunghwan, A., Jinwoo, C., Minyong, C., Wan Kyun, C. 2006. Metric slam in home environment with visual objects and sonar features. In *Proceedings of the 2006 IEEE International Conference on Intelligent Robots and Systems*. pp. 4048-4053.

[22] SungHwan, A., Jinwoo, C., Nakju Lett, D., Wan Kyun, C. 2008. A practical approach for EKF-SLAM in an indoor environment: fusing ultrasonic sensors and stereo camera. *Autonomous Robots*. Vol. 24, no. 3, pp. 315-335.

[23] Swain, M., Ballard, D. 1991. Color Indexing. *International Journal of Computer Vision*. 7:1, p 11-3.

[24] Tamimi, H., Andreasson, H., Treptow, A., Duckett, T., Zell, A. 2005. Localization of Mobile Robots with Omnidirectional Vision using Particle Filter and Iterative SIFT. In *Proceedings of the 2005 European Conference on Mobile Robots (ECMR05)*

[25] Tardif, J., Pavlidis, Y., Daniilidis, K. 2008. Monocular visual odometry in urban environments using an omnidirectional camera. *International Conference on Intelligent Robots and Systems*. pp. 2531-2538.

[26] Wolf, J., Burgard, W., Burkhardt, H. 2005. Robust vision-based localization by combining an image-retrieval system with Monte Carlo localization. *IEEE Transactions on Robotics*. Vol. 21, no.2, pp. 208-21