Sky segmentation with ultraviolet images can be used for navigation

Thomas Stone *
School of Informatics
University of Edinburgh
Edinburgh, EH8 9LE, UK
t.j.stone@sms.ed.ac.uk

Michael Mangan *
School of Informatics
University of Edinburgh
Edinburgh, EH8 9LE, UK
mmangan@staffmail.ed.ac.uk

Paul Ardin
School of Informatics
University of Edinburgh
Edinburgh, EH8 9LE, UK
s8528002@sms.ed.ac.uk

Barbara Webb
School of Informatics
University of Edinburgh
Edinburgh, EH8 9LE, UK
bwebb@inf.ed.ac.uk

Abstract—**Inspired by ant navigation, we explore a method for sky segmentation using ultraviolet (UV) light. A standard camera is adapted to allow collection of outdoor images containing light in the visible range, in UV only and in green only. Automatic segmentation of the sky region using UV only is significantly more accurate and far more consistent than visible wavelengths over a wide range of locations, times and weather conditions, and can be accomplished with a very low complexity algorithm. We apply this method to obtain compact binary (sky vs non-sky) images from panoramic UV images taken along a 2km route in an urban environment. Using either sequence SLAM or a visual compass on these images produces reliable localisation and orientation on a subsequent traversal of the route under different weather conditions.**

I. Introduction

The recent success of autonomous robotic vehicles [26] owes much to the availability of specific sensor technologies (LIDAR scanners, IMU, HD cameras, RADAR, GPS) and the continued development of computational resources capable of processing the data they generate. However, for the wider deployment of robots, particularly into price and energy critical markets, it is essential to bring down the hardware, computational, and energy costs of navigation. With this goal in mind, we take our inspiration from desert ants, which provide proof of principle that low-power and low-computation methods for navigation in natural environments are possible. Desert ants forage individually and without the use of chemical trails, yet can efficiently relocate their nest or a food source over long distances using visual cues alone [29]. The ant's compound eye provides very coarse resolution (on the order of 4° per pixel), and thus is unlikely to be capable of extracting and matching many distinct features, a capacity which is fundamental to many current visual Simultaneous localisation and mapping (SLAM) methods (e.g. [5]). Rather, their visual system appears fine-tuned to extract a key global feature – the panoramic skyline [16] – which behavioural observations have shown is sufficient for guidance in their complex visual habitat [7].

It has also been demonstrated that panoramic skylines are sufficient for robot localisation in urban environments [21, 14, 18]. The skyline (or more precisely, the horizon) has also been investigated as a visual cue for stabilisation in UAVs [19]. However, reliable visual segmentation of sky from terrain using conventional visual wavelength imaging appears difficult [18] particularly in changing weather conditions. Image parsing methods, e.g. [25], can be used to label sky amongst other regions, but the most successful current methods depend on advanced machine learning methods over large numbers of sample images to estimate the characteristics of sky. A range of approaches specific to detecting the sky region only are summarised in [20]. These include combining colour and texture characteristics, modelling the colour gradient, using connectedness or picture region constraints, and using alternative colour spaces and classification algorithms such as support vector machines. Shen and Wang [20] propose a greyscale method that uses gradient information and energy function maximisation, combined with several picture region assumptions, to obtain results on the order of 95-96% correct classification.

Far-infrared (IR) has been used to extract skylines, even after sundown [14], but is not robust for sky segmentation across weather conditions [4]. Clouds and sky appear very different in near-IR images (see supplementary material); indeed IR is often used explicitly for cloud detection [8].

By contrast, desert ants show sensitivity solely in the ultraviolet (UV) ($\sim 350nm$) and green ($\sim 510nm$) wavelengths [17, 11]. The effectiveness of these two wavelengths for sky detection was investigated in [16] using a custom sensor with one photodiode sensor for each. The results suggested using UV and green as opponent channels would be the most effective way to segment the sky under dynamic illumination conditions. In a follow up study [9] using five different wavelengths it was found that UV vs IR performed slightly better. However, because UV-imaging is limited by the filtering in standard camera optics, it has received little further investigation in robotics. In [3] a UV-receiving camera was paired with an omnidirectional mirror, but used to assess the directional information available in the polarisation of UV-light rather than to detect sky per se (sky segmentation by measuring the degree of polarisation of light is suggested in [19]). To our knowledge, there is only one recent investigation [23, 24] involving use of a UV-passing filter with a camera to enhance the difference between sky and ground regions, in the context of horizon determination for stabilisation. Neither the accuracy or consistency of UV vs visible wavelength imaging for detecting sky was reported in this work. UV has occasionally been used on robots for other purposes, including

* First author

object marking [10] and identification [28].

For proof of concept, we investigated the utility of UV vs visible light cues for sky segmentation using an adapted digital single-lens reflex (DSLR) camera system allowing wide spectrum imaging in high resolution. We demonstrate a simple thresholding algorithm applied to UV-only images is sufficient for accurate (compared to ground truth) and repeatable (comparing time-lapsed images across weather conditions) segmentation, and that this can be used to create panoramic segmented sky images that support navigation in an urban environment. The key features could be rapidly realised using off-the-shelf components (e.g. CCD without colour filters) providing a new class of low-cost outdoor navigation system.

II. METHODOLOGY

The aims of this study were: first, to assess the relative accuracy of sky segmentation using different wavelengths, over a range of different locations and conditions; second to assess the consistency of segmentation for UV vs visible wavelengths in the same location but at different times of day and weather conditions; and third to test the effectiveness of the segmented sky shape as a cue for robot navigation. Here we describe in detail the imaging device, the data sets collected for each test, and the segmentation and navigation methods used.

A. Full Spectrum Imaging

To sample images in both UV and visible spectra the internal cut filter of a standard DSLR camera (Nikon D600) was replaced by a fused quartz window (Advanced Camera Services Ltd, UK), and fitted with a high UV transmission wide angle lens (Novoflex Noflexar 3.5/35mm). Images could then be recorded with different filters: Thor Labs FGB37 bandpass filter (335 - 610 nm) to sample only visible light; Baader U-Filter (350 nm, half-width 50 nm) to approximate the ant's UV photoreceptor; and Knight Optical 520FCS5050 bandpass filter (515nm) to approximate the ant's green photoreceptor profile. The camera aperture was set to $f/16$, sensor sensitivity (ISO) to 200 and for each image an appropriate shutter speed was automatically selected to provide proper exposure given the chosen aperture. All images were saved using the uncompressed 14 bit Nikon Electronic Format (NEF) file format and subsequently demosaiced to a true colour 16 bit Tagged Image File Format (TIFF). Visible light images were stored in RGB format, the CMOS red channel only used for UV images, and a greyscale conversion of all three channels for green filtered images. To mimic the bichromaticity of the ant visual system, UV-green (UV-G) images were created by combining separate UV and Green intensity images into a 2 dimensional false colour space. All images were then scaled down by a factor of 10 along each dimension to give a resolution of 604×403 pixels prior to analysis.

B. Diverse Image Database

In July 2013, we took around 750 photographs in a variety of locations, including the natural habitat of desert ants (Seville, Spain) and wood ants (Sussex, UK) and a series of urban settings. UV-passing, green-passing and visible light-passing filters were used in turn. We selected a subset of 18 images representing a wide diversity of environments and conditions (see table I for example images) for which ground-truth sky segmentation was established. These included images expected to be difficult to segment due to surfaces reflecting the sky, such as water and windows. We used semi-automated labelling of sky and ground pixels through iterative applications of the GrowCut algorithm with user-correction [27]. We note that it was not possible to use existing 'benchmark' labelled datasets as these do not include UV information.

C. Timelapse Database

A time-lapsed image database was gathered from a single vantage point in January 2014 by mounting the camera on a secure platform and taking images across 4 days at 10 minute intervals in varying weather conditions (15:35-17:55 low light, 10:25-11:55 overcast, 10:50-11:00 sunny and 10:50-11:10 partial cloud). As before, all three filter types were used at each time point.

D. Urban Route Database

To allow the entire panorama to be recorded in a single image, the full spectrum camera was mounted above a panoramic mirror (Kugler, GmbH), fixed via a UV-passing quartz cylinder with 2mm wall (Fig. 1a), and images recorded using the UV-passing filter. A Sony Bloggie MHS-PM5, fitted with Bloggie 360 video lens kit attachment, was used to simultaneously capture panoramic images in the visible light spectrum. The combined system was mounted on a pole to elevate the horizon above a human carrier, with an approximate height of recording of 200 cm. Both cameras recorded video while the carrier walked a 2 km route that included narrow streets, city parks and squares, with moving traffic and pedestrians. True position was logged once per second using a GPS watch (Suunto Ambit). For testing navigation we used two videos of the route that were sampled on different days, at different times of year (January and April, introducing a change in the foliage), under different weather conditions (overcast and blue sky with clouds), and with different sun positions (11:40 and 16:47). Each video was converted to JPG images at 1 fps, providing two sets, consisting of 1337 and 1452 images respectively.

E. Sky Segmentation

To remove any distortion effects caused by camera movement when changing filters, images were aligned prior to segmentation by mutual information based image registration [13, 1]. K-means ($K = 2$, $r = 7$, sq. Euclidean distance) and Gaussian mixture model (GMM) clustering algorithms were implemented to assess both the accuracy and computational ease with which the sky could be automatically segmented from terrain. Pixels were clustered using image intensity only, ignoring any spatial connectivity information. These methods allowed comparison of performance despite the three image
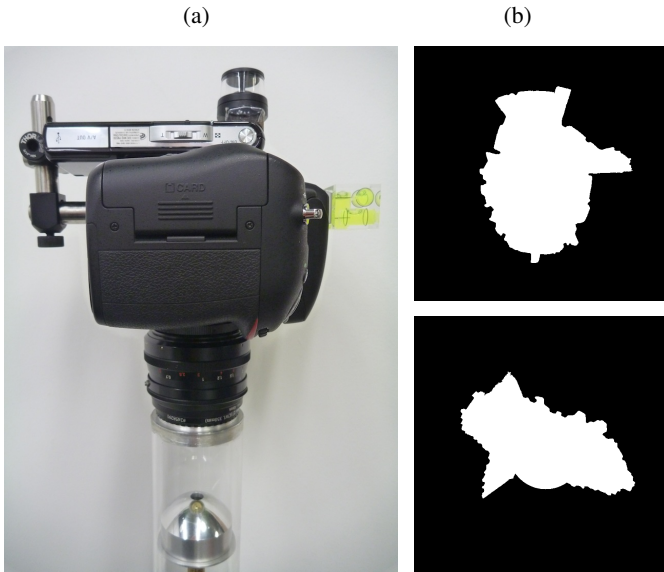
(a)   (b)

Fig. 1: **The panoramic camera set-up used for the urban route database.** (a) DSLR fitted with UV filter, quartz tube and parabolic mirror, with Sony Bloggie mounted on top. (b) Example of UV-segmented binary sky shape images used by our navigation algorithms. For each image the sky is shifted so that the centre of gravity of the shape is central in the image, to remove effects of tilt.

types having differing dimensions: UV=1D; UV-G=2D; and visible=3D (RGB).

We also tested a simpler segmentation algorithm tailored to the 1D UV intensity images. A histogram is computed from the intensity image, excluding the darkest and brightest 2 percent of pixels, and the adapted-watershed algorithm described in [30] is applied until only two segments of the histogram remain. The intensity value midway between the closest points of the two histogram segments is then set as the threshold for ground/sky segmentation. This method is robust to multiple maxima caused by a visible sun and ground reflectance, which were problematic for alternative algorithms.

F. Navigation Tests

OpenSeqSLAM [22], an open source Matlab implementation of SeqSLAM [15] , was used to test whether panoramic UV-segmented images contain sufficient information for localisation. Images were preprocessed by applying a mask to disregard sectors of the image that were not parts of the panorama in the parabolic mirror. The simple watershed segmentation algorithm described above was then applied to produce 90×90 binary images (sky vs non-sky). The sky region was isolated by boundary tracing from the centre of the image, and shifted by centralising the centre of gravity to reduce effects of tilt (Fig. 1b). Performance of our UV-segmented image set was compared to regular SeqSLAM using a set of visible light panoramic images recorded at the same locations and with a similar number of bits per image (16×16, 8 bit greyscale).

We tested localisation on the April route with the January reference images and vice versa. To distribute effects caused by the two sequences being particularly in or out of sync, we compared a reference set starting at the first frame of the route and sampled every 10 seconds against 10 test sets, starting with the first, second, third etc. frame of the other route and sampled every 10 seconds thereafter. SeqSLAM was run in both directions for each of the resulting 19 route pairs. SeqSLAM first performs an image difference function on the current image and all training images, stores the results in a matrix, and uses contrast enhancement to find the local best matches. A familiar sequence is then found by scoring trajectories across the matrix to look for those with the combined best match. The trajectory with the lowest score is considered a match if it is significantly smaller than all other scores outside a sliding window ($R_{window} = 10$).

We also tested an alternative navigation method that has been suggested as a possible mechanism used by ants, the visual compass algorithm [6, 12]. The basic assumption of this algorithm is that rotating a test image until the difference from a training image is minimised will indicate an approximately correct heading, under the assumption that captured training images are generally oriented in the direction of forward travel along the route. Although some versions of this algorithm require knowledge of the sequence to constrain which images to compare [12] it has also been shown to work effectively by taking the minimum over all rotations over all training images [2], i.e., assuming no sequence memory. This implicitly also returns a location estimate: the training image most similar to the test image after rotation.

Using the extracted skyline image sets at full resolution (691×691 pixels) a sub-sample of 1 image every 10 seconds from the test set was compared at 90 4° rotations against every image in the full 1fps training set, by summing the output of an *XOR* function. The location and rotation producing the smallest overall difference was chosen as best match. A comparison was made to both greyscale and colour visible light images, unwrapped and scaled to a similar number of bits; 180×28 and 90×14 pixels respectively.

III. RESULTS

A. Automatic Segmentation of Sky is Best Using UV Intensity Images

Table I shows the results of automatic sky segmentation, compared to manually labelled ground-truth, in the first image-database. Twelve example images are shown; the mean accuracy is based on 18 images. Segmentation using K-means clustering is most accurate using UV intensity images, with a mean accuracy of 96.8%. This is significantly better than visible light (mean accuracy 89.9%, paired Wilcox test $p = 0.0018$) and UV-G (mean accuracy 93.2%, paired Wilcox text $p = 0.025$). The accuracy for UV is also significantly less variable (standard deviation 3.6) than either visible light (s.d. 12.3, F-ratio test $p < .0001$) or UV-G (s.d. 9.5, F-ratio test $p = 0.0002$). Segmentation in the visible spectrum was not

TABLE I: **Accuracy of sky segmentation across locations.** A subset of our hand labelled image database is shown below to illustrate the diversity of images collected. Image triplets (Visible, UV-G and UV) were sampled locations chosen to assess the performance of sky segmentation in diverse surroundings. The percentage of correctly labelled pixels, when compared to a user defined ground truth, is shown for each image set. Mean scores (final row) are for the full database. Overall, for clustering methods performance is best using k-means on UV images, however similar results can be achieved with the computationally simpler adapted-watershed algorithm.

| Image	Feature	K-means, $O(kDN)$			GMM, $O(iND^2)$			Watershed, $O(N)$
		Visible	UV	UVG	Visible	UV	UVG	UV
	Desert scrub, Seville	80.283	89.102	89.035	76.742	82.745	78.141	**91.107**
	Forest floor, Sussex	92.251	93.058	86.255	71.17	83.814	72.851	**93.28**
	Pond reflecting sky	93.245	97.15	95.459	95.63	92.911	90.489	**93.686**
	Landscape with dark clouds	52.42	97.731	59.771	**99.058**	98.198	58.491	98.563
	Building hidden by trees	**98.137**	97.817	98.126	91.452	94.665	92.284	97.689
	Minaret with dark clouds	75.549	97.75	90.272	**98.796**	98.155	98.538	98.511
	Minaret with white clouds	75.889	99.001	97.083	71.201	99.076	**99.171**	99.084
	White building on overcast day	**99.53**	99.353	99.336	91.805	91.334	99.33	99.512
	City street with grey sky	99.047	99.123	98.987	98.865	98.878	96.581	**99.263**
	Building with metallic roof	85.348	86.32	**87.298**	76.967	84.599	81.364	84.687
	Building with metal wall	96.366	98.595	98.218	61.211	**98.722**	91.661	98.901
	Buildings with reflective windows	97.612	**98.068**	97.662	71.548	81.109	76.985	97.754
	Mean accuracy (n=18)	89.867	96.778	93.229	85.962	92.36	87.448	**96.939**
	std. dev.	12.245	**3.631**	9.497	12.021	7.514	11.682	3.883
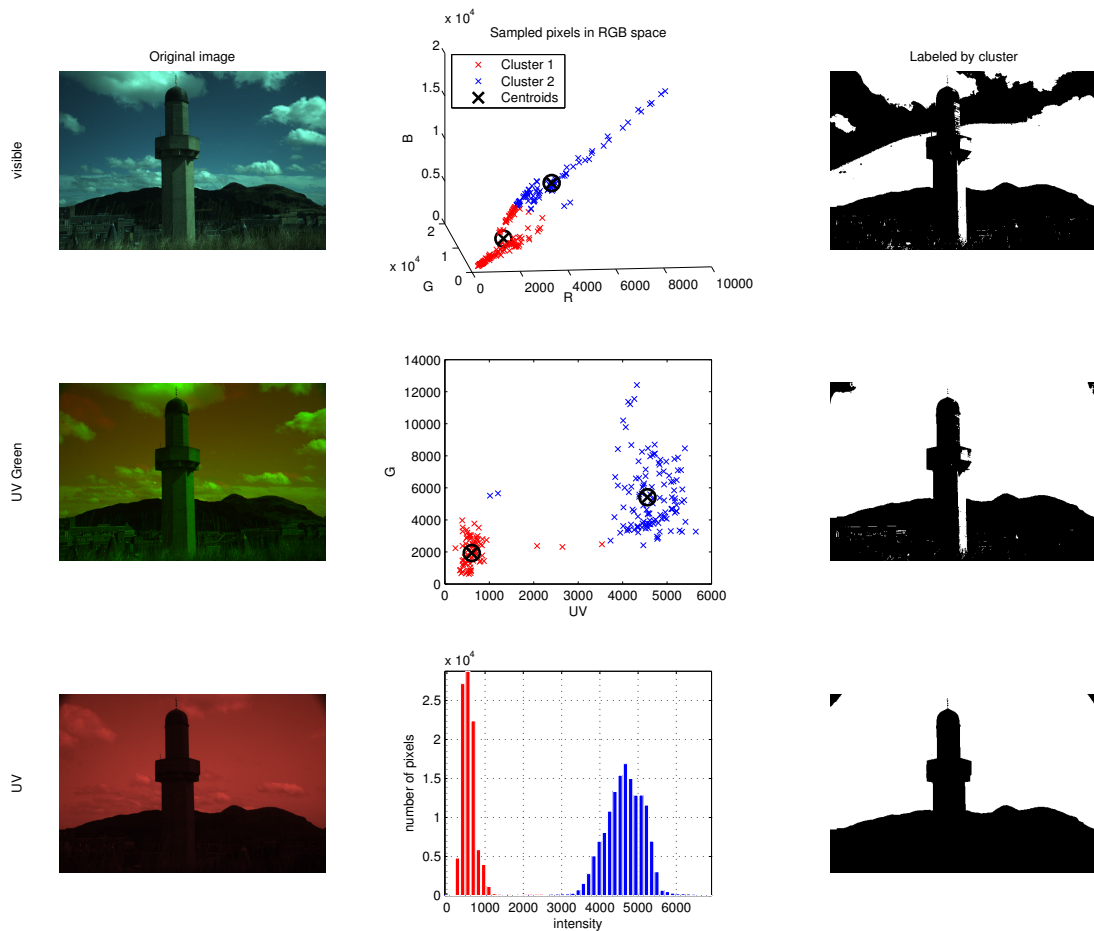
Fig. 2: **Typical example of conditions in which UV-intensity segmentation outperforms both visible and UV-G images. Performance is worst using visible light as sky and ground are not easily separable when clouds are present. In contrast, sky and ground naturally separate along the UV dimension giving robust segmentation in UV and UV-G images.** Raw images are shown in the left column. The central column shows the spread of image pixels plotted in their respective colour space, coloured with respect to the attempted automatic k-means segmentation (blue=sky, red=ground). In each scatter plot a random subsample of 200 pixel values was used. The right column shows the resulting segmented images with white representing sky and black representing ground. Note that the black marks in the corners of the the UV pictures are the result of vignetting due to an overlapping lens hood that is part of the Baader U-filter.

improved using the more robust Gaussian mixture model classifier (mean accuracy 86.0%). Visible light sky segmentation was also tested in other colour spaces, such as the A and B dimensions of LAB space, however no notable improvements were observed over RGB for any of our clustering algorithms. The simple watershed algorithm for UV performs just as well as K-means clustering (mean accuracy 96.9%) but at less computational cost ($O(N)$ instead of $O(kDN)$), where k is the number of clusters, D the dimensionality of our pixel space and N the number of pixels). The level of accuracy obtained is comparable with that reported in other state of the art sky segmentation methods, e.g. 96.05% reported in [20]. In fact a proportion of our remaining error was simply due to vignetting of the overlapping lens hood (see Fig. 2 middle row). We also note that the 'ground truth' labelling was performed on the visible images and so was if anything biased against UV.

Fig. 2 shows the underlying reasons for better performance of UV. In the visible light image (top left), the bright white clouds form a cluster. As the classifier is constrained to two groups, the blue sky is incorrectly labelled the same as the ground. This is the most common cause of error when using visible light, but reflectance from buildings and sun flare also cause segmentation irregularities. In contrast, in the UV-G image the clouds are labelled with the sky, and only reflection from the buildings causes some segmentation issues. The pixel-plot shows that the successful separation is attributable mostly to the UV channel, and in fact the segmentation is improved by using the UV channel only, which does not classify strong lighting on the buildings as sky.

The second image-database allows us to test both the accuracy (vs ground truth), and the reliability (vs images from the same location) of sky segmentation across weather and

lighting conditions. This is important as it may be acceptable for a navigating robot to misclassify some portion of the ground as sky (or vice versa) provided the classification remains consistent over time. Accuracy is compared against user labelled ground truth, and here again we find that segmentation is more accurate using UV intensity images (mean=99.5%, s.d=0.084) than visible (mean=95.2%, s.d.=5.327). Reliability was assessed by measuring the mean entropy and Pearson correlation coefficients of the binary labelled images. Specifically, the mean entropy per pixel location, $H_2(x_i)$, in labelled UV and visible images was calculated as

$$p(x_i) \approx \frac{\sum_{t=1}^{T}(x_{it})}{T} \quad (1)$$

$$H_2(x_i) = -p(x_i)\log_2 p(x_i) - (1-p(x_i))\log_2(1-p(x_i)), \quad (2)$$

where x_{it} is the intensity value of a pixel at location i at time t, and T is the number of images in the set. The Pearson correlation is given by

$$r(\mathbf{x},\mathbf{y}) = \frac{cov^*(\mathbf{x},\mathbf{y})}{var^*(\mathbf{x})var^*(\mathbf{y})}, \quad (3)$$

where $cov^*(\cdot,\cdot)$ indicates the sample covariance between two images $\mathbf{x}$ and $\mathbf{y}$ and $var^*(\cdot)$ the sample variance of an image. Fig. 3 presents the resultant entropy and correlation plots for visible and UV image sets. It is clear that pixels in the visible domain are highly variable (mean entropy=0.1877) particularly in the crucial area where buildings meet the sky. There is also significant interference from sunlight on buildings and street lights. This variance results in low correlation coefficients between images throughout the data-set. Labelled images that correlate poorly in the heat map in Fig. 3c (images 5, 15 and 20) corresponded to the 3 images shown in the left column of the Fig. 3a, where cloud, sunlit buildings and low lighting caused discrepancies in segmentation. The causes of these problems are clearly visualised by viewing the plot showing average entropy per pixel (Fig. 3b). In contrast, UV images have a low mean entropy per pixel (0.0048), producing a highly correlated and thus reliable dataset across conditions, even at low light levels.

B. UV Sky Segmented Images Are Sufficient For Navigating Cities

A set of images recorded along a route can be used in several ways for subsequent navigation. We first tested whether binary sky/non-sky images derived from panoramic UV images recorded along a 2km urban route (see supplementary video) could be used for localisation (or loop-closing) using SeqSLAM as described in section II-F. Overall, using a window length of 10 images, we can obtain a matching precision of 100% (no false positives) with a recall of 81%; at 100% recall, the precision is 94%; these values compare well to previous results for this algorithm [15]. Fig. 4a shows where matching is unsuccessful, which appears to correspond to locations under trees, where the sky is either obscured or its
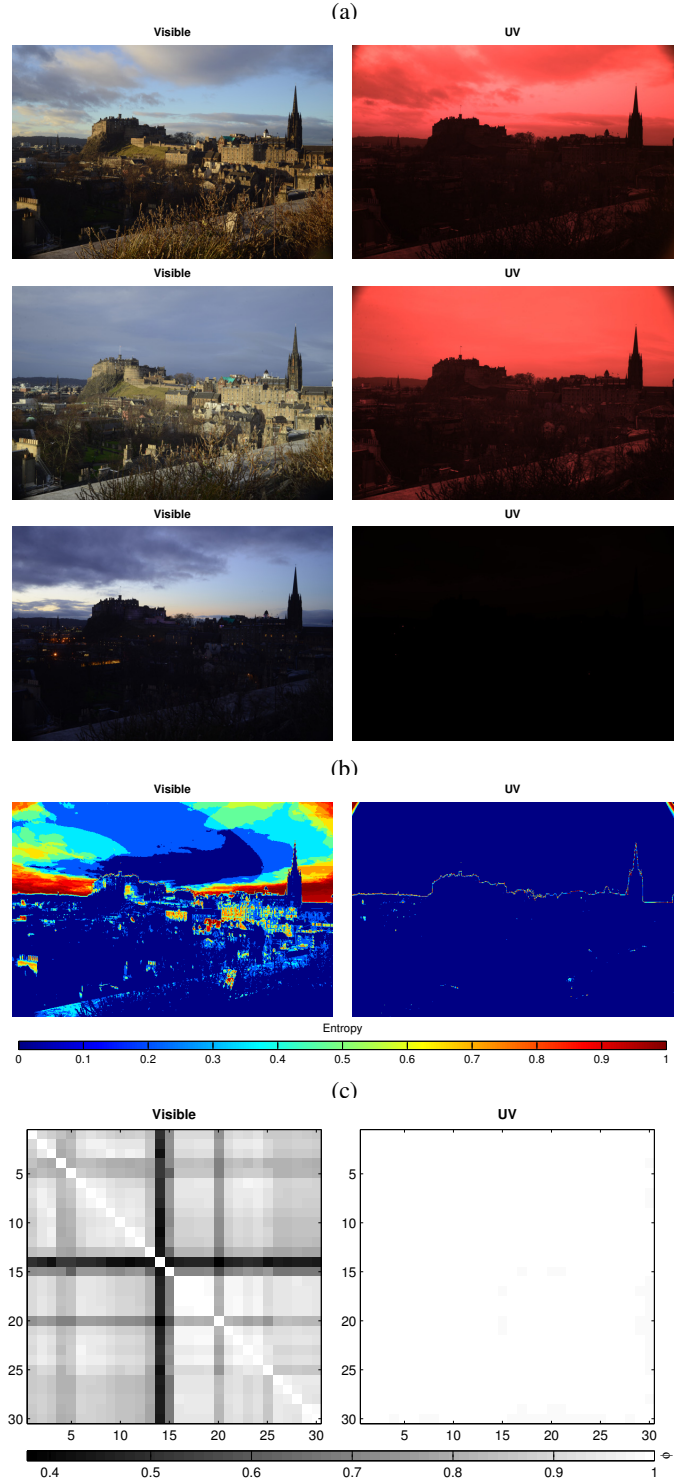

Fig. 3: **Reliability of sky segmentation across weather conditions.** (a) Time-lapse images sampled under differing weather conditions (note that the UV pixel intensities have been increased $5\times$ to make image more visible, but are still extremely low in the third example). (b) Pixel-wise entropy and (c) Pearson correlation coefficients computed across the image dataset. Both indicate high variance in visible light images but consistent labelling using UV images.

segmented shape changing too rapidly for the sparse temporal sampling used here. Our result for binary UV images is only a few percent below that obtained using standard SeqSLAM on the equivalent panoramic greyscale visible light images (precision 97%), which deals better with situations where there is little or rapidly changing sky shape information (under trees) as it can still exploit information in the ground region. On the other hand standard SeqSLAM tended to fail in areas where the ground region is relatively featureless but there are substantial changes in the sky region (clouds) between training and test sets.

While SeqSLAM is highly effective for localisation if following the same route, it does not provide a mechanism to stay on the route. Hence we also test the visual compass algorithm, as described in II-F which produces both a location match and an estimate of the heading direction needed to stay on route. The obtained precision using binary UV images was 84.7%. Fig. 4c shows the heading direction suggested by the visual compass at each point on the route; it can be seen that these align well with the required direction to stay on the route. Using the equivalent visible light images produced a significantly worse performance with a precision of only 44.1% and many heading direction choices that would lead away from the route (Fig. 4d) .

IV. Conclusions

It has been suggested that UV light sensing in desert ants is a specialisation that supports their impressive navigation by enhancing sky-ground contrast. We have demonstrated here that UV images are highly effective for sky segmentation, that the segmented sky shape is robust to changing light and weather conditions, and that the resulting binary sky/non-sky images can support effective navigation.

For a wide range of locations and light conditions, automated sky segmentation is significantly more effective for UV than visible light. Although only a limited set of images were hand-segmented for ground truth evaluation, visual inspection of UV segmentation in our larger collection of images suggests broad reliability, in particular, robustness against all kinds of cloud cover. The unidimensional signal also allows the application of a computationally cheap segmentation algorithm that is just as effective as K-means clustering.

There were a few cases where ground objects, such as particular kinds of metal roofing, had high UV reflectivity. However, for navigation, it is less important to find the skyline exactly than it is to find it consistently. By testing sky-ground segmentation for images taken from the same location on different days, at different times of day, and under different weather conditions, we found that UV images have an extremely high consistency in pixel classification, around 40 times better than visible light.

The navigation potential of a panoramic UV camera system was demonstrated by recording 2km routes in an urban environment on two different days with different weather conditions. Using either SeqSLAM or a visual compass, matching of UV-segmented binary images of the sky contour from one day against another supported successful localisation. For the visual compass a reliable indication of the correct heading direction to continue along the route was also obtained. Using non-segmented visible light images of a similar file size, SeqSLAM was still effective for localisation, but determining location or heading direction from the visual compass was far less effective. Failure of the visual compass is generally due to mismatch or aliasing, i.e., a training image far from the current location provides, at some arbitrary rotation, a closer match than the nearest image. This suggests its effectiveness could be improved (and its application sped up) by using SeqSLAM as a preprocessing step to find the best match, then finding the minimum over image rotations for this image only.

We note that the skyline is a particularly consistent aspect of urban (and some other) environments. Many changing features such as traffic and pedestrians, fall entirely below the skyline, and different lighting conditions such as sun falling on one or the other side of street make little difference to the segmented shape. Weather changes occur above the skyline. Consequently it appears that a substantial amount of information is retained by reducing images to binary sky vs non-sky: SeqSLAM performs almost as well on such images as on full grey scale images. The few situations in which the system failed were in locations where very little sky was visible (under scaffolding, or under trees). The considerable difference in foliage between our two datasets, and the 10 second time intervals between frames could cause a notable difference in sky shape directly above the camera between training and test sets in this situation. This could perhaps be rectified by smoothing over recent frames to filter out fast moving shapes. Alternatively, rather than using binary images, the UV threshold could be used to mask out the sky region (including uninformative features such as clouds) but information below the skyline, possibly from visible light channels, could be retained to improve robustness.

There are a number of potential applications for this approach. It could be used effectively in urban environments where GPS can be unreliable. The nature of the segmented image also makes it trivial to perform corrections to tilt by simply offsetting the image depending on the centre of gravity of the sky pixels. It could therefore also be potentially useful for robots in rugged environments, where SLAM needs to be reliably carried out on uneven terrain. Sensitivity to UV light at low light levels could be improved by using a CCD chip without colour filters, rather than the converted visible light camera used in these tests. The relatively simple computation to segment sky could be implemented in hardware, providing a compact and low-power sky detection sensor, which could offer cheap high precision localisation in many outdoor autonomous robot applications.
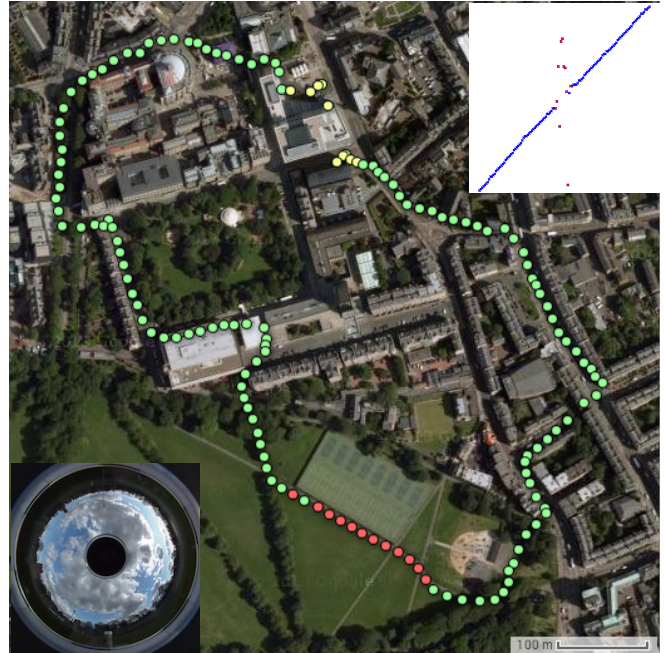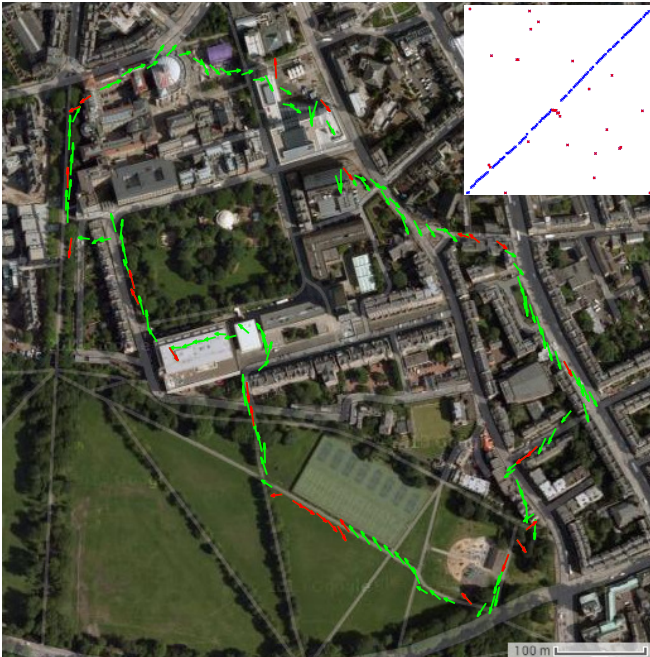
(a) UV sky segmented SeqSLAM
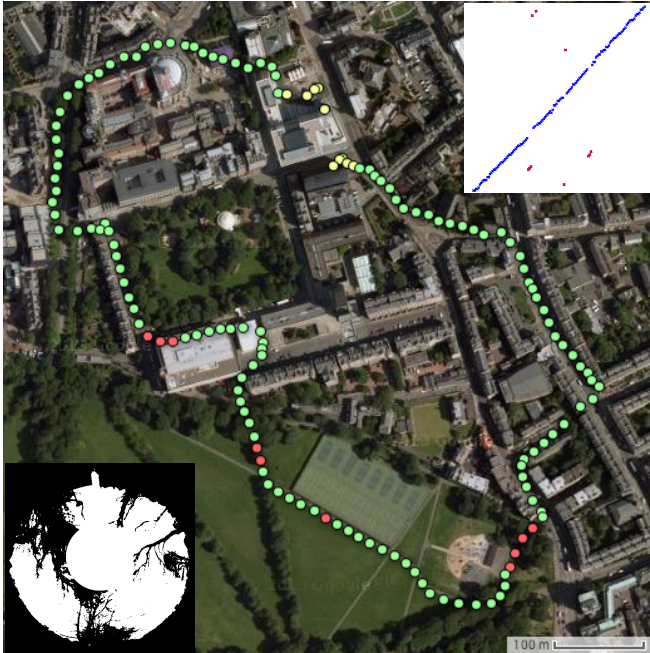
(b) Visible light full image SeqSLAM

(c) UV sky segmented Visual Compass

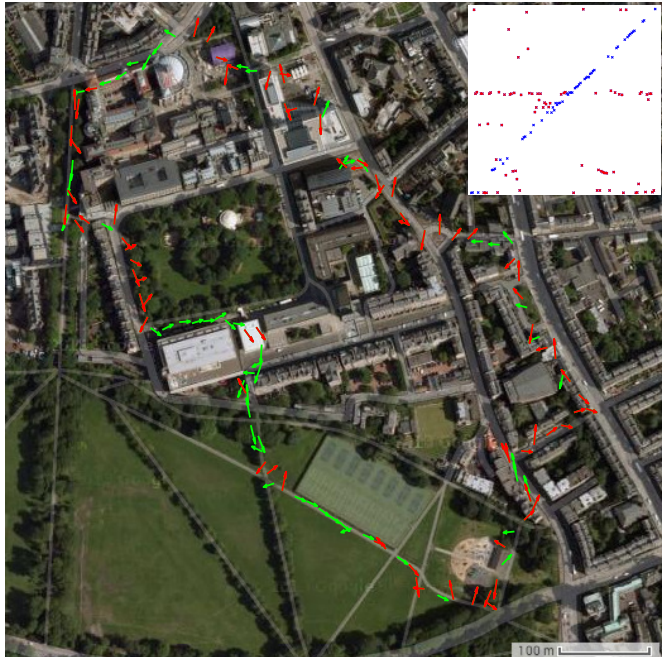(d) Visible light full image Visual Compass

Fig. 4: **Using sky-segmented images to navigate a city.** (a) and (b) plot the example results of the SeqSLAM localisation trial taken from the worst test set for UV (left) and the worst for visible light (right). Green dots represent locations that were correctly recognised and red dots those that were not. Yellow dots are at the beginning and end of the route, where the sequence was too short to localise. In the top right corner of these figures the estimated location index of the training set is plotted against the location indices of this test set. In the bottom left corner an example panoramic image is shown, corresponding to an incorrectly localised image in this test set. Branches (for UV) and clouds (for visible light) were typical features that caused the algorithm to fail.

(c) and (d) show the heading direction that the visual compass algorithm would select in order to retrace the route (left for UV, right for visible light), with all headings corresponding to a correctly matched location coloured green. Performance is better for UV.

Map data: Google, Infoterra Ltd & Bluesky.

REFERENCES

[1] Artyushkova, K. Automatic Image Registration using (Normalized) Mutual Information for users of IP toolbox, 2006. URL http://www.R-project.org/.

[2] B. Baddeley, P. Graham, P. Husbands, and A. Philippides. A model of ant route navigation driven by scene familiarity. *PLoS Computational Biology*, 8(1):e1002336, January 2012. doi: 10.1371/journal.pcbi.1002336.

[3] N. Carey and W. Stürzl. An insect-inspired omni-directional vision system including UV-sensitivity and polarisation. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 312–319. IEEE, 2011. doi: 10.1109/ICCVW.2011.6130258.

[4] L. Chapman, J.E. Thornes, J-P. Muller, and S. McMuldroch. Potential applications of thermal fisheye imagery in urban environments. *Geoscience and Remote Sensing Letters, IEEE*, 4(1):56–59, 2007. doi: 10.1109/LGRS.2006.885890.

[5] M. Cummins and P. Newman. FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance. *The International Journal of Robotics Research*, 27(6):647–665, June 2008. doi: 10.1177/0278364908090961.

[6] M.O. Franz, B. Schölkopf, H.A. Mallot, and H.H. Bülthoff. Where did I take that snapshot? Scene-based homing by image matching. *Biological Cybernetics*, 79(3):191–202, 1998. doi: 10.1007/s004220050470.

[7] P. Graham and K. Cheng. Ants use the panoramic skyline as a visual cue during navigation. *Current Biology*, 19(20):R935–7, November 2009. doi: 10.1016/j.cub.2009.08.015.

[8] D.I. Klebe, R.D. Blatherwick, and V.R. Morris. Ground-based all-sky mid-infrared and visible imagery for purposes of characterizing cloud properties. *Atmospheric Measurement Techniques*, 7(2):637–645, 2014. doi: 10.5194/amt-7-637-2014.

[9] T. Kollmeier, F. Röben, W. Schenck, and R. Möller. Spectral contrasts for landmark navigation. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 24(1):1–10, January 2007. doi: 10.1364/JOSAA.24.000001.

[10] S. Komai, T. Kuroda, and M. Takano. Development of Invisible Mark and Its Application to a Home Robot. In *Service Robotics and Mechatronics*, pages 171–176. 2010. doi: 10.1007/978-1-84882-694-6_30.

[11] T. Labhart. The electrophysiology of photoreceptors in different eye regions of the desert ant, Cataglyphis bicolor. *Journal of Comparative Physiology A*, 158(1):1–7, 1986. doi: 10.1007/BF00614514.

[12] F. Labrosse. The Visual Compass : Performance and Limitations of an Appearance-Based Method. *Journal of Field Robotics*, 23(10):913–941, 2006. doi: 10.1002/rob.

[13] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE transactions on medical imaging*, 16(2):187–98, April 1997. doi: 10.1109/42.563664.

[14] J. Meguro, T. Murata, Y. Amano, T. Hasizume, and J. Takiguchi. Development of a Positioning Technique for an Urban Area Using Omnidirectional Infrared Camera and Aerial Survey Data. *Advanced Robotics*, 22(6-7):731–747, January 2008. doi: 10.1163/156855308X305290.

[15] Michael J Milford and Gordon Fraser Wyeth. Seqslam: Visual route-based navigation for sunny summer days and stormy winter nights. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 1643–1649. IEEE, 2012.

[16] R. Möller. Insects could exploit UV-green contrast for landmark navigation. *Journal of theoretical biology*, 214(4):619–31, February 2002. doi: 10.1006/jtbi.2001.2484.

[17] M.I. Mote and R. Wehner. Functional characteristics of photoreceptors in the compound eye and ocellus of the desert ant, Cataglyphis bicolor. *Journal of comparative physiology*, 137(1):63–71, 1980. doi: 10.1007/BF00656918.

[18] S. Ramalingam, S. Bouaziz, P. Sturm, and M. Brand. Geolocalization using skylines from omni-images. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 23–30. IEEE, 2009. doi: 10.1109/ICCVW.2009.5457723.

[19] A.E.R. Shabayek, C. Demonceaux, O. Morel, and D. Fofi. Vision Based UAV Attitude Estimation: Progress and Insights. *Journal of Intelligent & Robotic Systems*, 65(1-4):295–308, 2012. doi: 10.1007/s10846-011-9588-y.

[20] Y. Shen and Q. Wang. Sky Region Detection in a Single Image for Autonomous Ground Robot Navigation. *International Journal of Advanced Robotic Systems*, 10:362, 2013. doi: 10.5772/56884.

[21] F. Stein and G. Medioni. Map-based localization using the panoramic horizon. *IEEE Transactions on Robotics and Automation*, 11(6):892–896, 1995. doi: 10.1109/70.478436.

[22] N. Sünderhauf, P. Neubert, and P. Protzel. Are We There Yet? Challenging SeqSLAM on a 3000 km Journey Across All Four Seasons. In *Proc. of Workshop on Long-Term Autonomy, IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, Germany.*, 2013.

[23] M.H. Tehrani, M.A. Garratt, and S. Anavatti. Gyroscope offset estimation using panoramic vision-based attitude estimation and Extended Kalman Filter. In *CCCA12*, pages 1–5. IEEE, December 2012. doi: 10.1109/CCCA.2012.6417863.

[24] Mohsen H. Tehrani, M. A. Garratt, and S. Anavatti. Horizon-based attitude estimation from a panoramic vision sensor. In Ghose Debasish, editor, *IFAC-EGNCA 2012*, pages 185–188, February 2012. doi: 10.3182/20120213-3-IN-4034.00035.

[25] J. Tighe and S. Lazebnik. Superparsing. *International Journal of Computer Vision*, 101(2):329–349, October

2012. doi: 10.1007/s11263-012-0574-z.

[26] Urmson, C. The self-driving car logs more miles on new wheels, August 2012. URL http://googleblog.blogspot.hu/2012/08/the-self-driving-car-logs-more-miles-on.html.

[27] V. Vezhnevets and V. Konouchine. GrowCut: Interactive multi-label ND image segmentation by cellular automata. In *Proc. of Graphicon*, pages 150–156, 2005.

[28] H. Wajima, T. Makino, Y. Takahama, T. Sugiyama, and A. Keiichi. An Autonomous Inspection Robot System for Runways Using an Ultraviolet Image Sensor. In *Intelligent Autonomous Systems 7*, pages 357–364, 2002.

[29] R. Wehner. The architecture of the desert ant's navigational toolkit (Hymenoptera: Formicidae). *Myrmecol News*, 12:85–96, 2008.

[30] N.S. Zghal and D.S. Masmoudi. Improving watershed algorithm with a histogram driven methodology and implementation of the system on a virtex 5 platform. *International Journal of Computer Applications*, 9(11): 29–35, 2010. doi: 10.5120/1435-1934.