

## Airborne Inspection using Single-Camera Interleaved Imagery

Michael Maurer, Andreas Wendel, Horst Bischof  
Institute for Computer Graphics and Vision  
Graz University of Technology, Austria  
{maurer, wendel, bischof}@icg.tugraz.at

**Abstract.** We propose an airborne inspection system which operates with a single camera. It provides live high-resolution detail images at low frame-rates for inspection, interleaved in time with several coarse-resolution overview images suitable for visual servoing. By defining the object of interest in 3D, the same pose estimation routine can be used for visual servoing and object tracking, which saves computation effort. Localization in 3D allows inspection from different perspectives without manual redefinition of the object of interest. Additionally, small accumulated localization errors of the image acquisition pipeline are corrected using correlation-based image stabilization. In our experiments, we show that the proposed method results in better image quality and higher frame-rates than the straight-forward approach of streaming wide field-of-view high-resolution images. Additionally, we evaluate the timings of our interleaved image acquisition. Given the requirement of one detail image per second, we can capture a maximum of 6 additional overview images per second. This utilizes the available on-board processing power and transmission bandwidth best and is sufficient for simultaneous inspection and visual servoing.

### 1. Introduction

Recently mobile image acquisition platforms such as micro aerial vehicles (MAVs) have become affordable and suitable for various inspection tasks, such as power pylon inspection. Usually the operator is skilled in visual inspection and not in flying an airborne vehicle, so the MAV should fly autonomously. Coarse position-hold functionalities are already available for outdoor environments using GPS; however, a disadvantage of autonomous navigation using GPS is drifting over time and the



Figure 1. Our inspection system is capable of providing wide field-of-view low-resolution overview images (left) and high-resolution detail images with a narrow field of view (right) using a single industrial camera mounted on an MAV. The inspector can interact with the tablet computer to select an arbitrary object of interest.

”blind” navigation. In other words, repeated inspection tasks have to be planned over and over again to ensure accurate positioning and collision avoidance. This recurrent planning process can be prevented by using vision-based approaches to MAV localization and positioning [15, 1], also called visual servoing.

The requirements on the camera for inspection compared to visual servoing differ considerably. For visual servoing a wide field-of-view is of advantage but in contrast for inspection a close-up view results in more accurate inspections. Full-frame, high-resolution imagery would enable both tasks to be performed with a single camera, but on current MAVs there is neither enough on-board processing power to perform visual navigation with high-resolution images nor enough bandwidth to stream high-resolution full-frames at appropriate frame rates to the ground. Therefore, two different types of cameras would be required, but this is not feasible given the limits on payload.

In this work, we present a method which allows to fulfill both requirements using a single camera by interleaving overview and detail images in time. We permanently switch the camera settings to gain on the one hand gray-scale low-resolution images with a wide field of view, which exploit the binning functionality of the camera to reduce motion blur and can be obtained at a relatively high frame rate. These *overview* images are well suited for visual servoing. On the other hand, we capture detailed colored inspection images at full resolution by setting an area of interest (AOI), which we call *detail* images. This leads to a massive increase of imagery that can be streamed over a wireless 802.11n link. On the ground, the inspector gets a live full-area overview and high-quality detail images visualized on a tablet computer. He can directly interact with the system to define the object he wants to inspect (see Figure 1).

## 2. Related Work

Using a mobile airborne platform for inspection tasks is beneficial for a human operator, as he obtains images from positions which are typically out of range. Supporting the inspector during his task is essential and requires three major components. First, a visual servoing approach is necessary to position the MAV at a certain point of interest in space. In turn, this requires accurate localization and mapping to run in real-time on the system or on the ground. Second, the object of interest should be tracked in the video sequence. Typical inspection focuses on a certain area of interest in the image, so the third component is the digital stabilization of the extracted image patch to ensure the possibility of high-quality inspection. Bottlenecks of such a system are the on-board processing power and the limited transmission bandwidth to the ground-station or mobile visualization client.

In the following paragraphs we present a selection of previously proposed inspection setups. Additionally, we introduce a number of algorithms used to overcome the limited transmission bandwidth, to perform visual localization and servoing, and for digital image stabilization.

### 2.1. Inspection Systems

Most work on visual inspection approaches based on MAVs focuses on powerline inspection. An overview of various systems has been presented by Katrasnik et al. [8]. The authors address inspection

systems using automated helicopters, flying robots, and climbing robots. The systems are assessed according to design requirements, inspection quality, autonomy and universality of the inspection.

Wang et al. [14] designed an inspection robot called SmartCopter. The system consists of two separate parts, an unmanned autonomous helicopter and an inspection system. The system is equipped with two inspection cameras, one for visible light and one for infrared imaging. Weng et al. [17] presented a VTOL (vertical take off and landing) flying robot for inspection and surveillance operations. The setup contains one wireless CCTV camera for inspection and is controlled by a pilot who is assisted by a stabilization system using accelerometer, compass, and gyro sensor data to stabilize the VTOL robot. Both inspection systems are not applicable for our proposed system because they require next to the inspector an additional pilot. Further they only provide inspection images and are therefore not extendable to autonomous visual navigation.

### 2.2. Image Encoding

An image contains a lot of information and thus encoding is necessary to reduce the required bandwidth for transmission, especially when wireless link are used.

Meuel et al. [10] developed a video codec to transmit full HDTV ( $1920 \times 1080$  px) resolution over a wireless transmission channel with a bit rate of 1-3 Mbit/s. They exploit the ROI detection of a standard AVC video codec and additionally use global motion compensation. They claim to preserve more details than an AVC codec at the same bit rate. As shown in their experiments, the codec is only suitable for high-altitude aerial images, which is not the case in our inspection system. Furthermore, a special encoder would require additional processing power on-board the MAV, which that is already a bottleneck of our system.

Zheng et al. [18] presented an improved region of interest (ROI) progressive image transmission algorithm. They first transmit the ROI coefficients and the wavelet coefficients important for human vision, followed by the background that is constrained by its bit rate using an expansion factor. This expansion factor gives the ability to control the image transmission. The approach would be suitable for transmitting inspection images containing very small objects of interest. However, in our system the object of in-

terest should be transmitted in high-resolution, so the ROI size does not vary significantly from the total image size.

Salous et al. [12] presented a method to prioritize spatial regions for transmission and therefore preserve visually informative data regions. They implement a context-based transmission of the image and start with an iconic image with coarse structures. By exploiting the knowledge about image composition, they segment it and prioritize single regions. The method was tested using neurological images and showed a reduction of the latency compared to progressive image transport, as well as a reduction of the transmission time of necessary regions to half the time compared to the full image. Exploiting the advantages of context-based transmission would be feasible for specific inspection approaches. In our system the object to track is arbitrary and therefore the context is a-priori unknown. Further, the on-board computing resources are limited and do not allow an additional segmentation of the image.

Finally, Gong et al. [5] proposed an image compressing method for unmanned aerial vehicles. They cluster the image sequence using the inertial navigation system and predict the position of the intermediate frames of one cluster using homography matching. Finally they merge the first frame and the transformed intermediate frames into a big image and stream the result using the JPEG2000 codec. This approach is suitable to get single, large-scale aerial images, but it is not applicable for inspection of single objects. Considering the usage of this approach would lead to one big image of the object. All acquired images would be clustered to one cluster because of the almost identical inertial navigation data while hovering.

### 2.3. Visual Localization and Servoing

The localization of the MAV in an unknown environment is addressed by a simultaneously localization and mapping (SLAM) approach. Several different methods for visual SLAM exist [3, 4]; a recent one which is widely seen as state-of-the-art is Parallel Tracking and Mapping (PTAM) [9]. In this approach, the mapping is performed using key-frames only and the tracking uses sparse feature points. It is suitable to run at low-performance processing units because of the separation of tracking and mapping in two decoupled tasks and has proven to work well in a visual servoing loop [1, 16]. A typical approach is to

employ on-board position control using a monocular camera as exteroceptive sensor and filter the results with data from the internal sensors, including an inertial measurement unit (IMU) and GPS.

Another way to localize the MAV in a previously reconstructed environment has been introduced by Wendel et al. [15]. Their localization approach scales well for outdoor environments due to the usage of virtual cameras. The output of the localization can again be used for visual servoing as described in [16].

### 2.4. Image Stabilization

Image stabilization is widely used in consumer cameras. On the one hand it can be achieved by optical stabilizers, on the other hand by image post processing, but satisfactory stabilization is only possible when using both methods.

Hsu et al. [7] presented a digital image stabilization (DIS) technique to remove the unwanted shaking phenomena of image sequences of hand-held cameras. To preserve intentional motions they use a motion estimation unit and an inverse triangle method to gain reliable motion vectors. Auberger et al. [2] presented a very low-cost and low-power video stabilization algorithm based on binary motion estimation in key areas of the image. After an outlier removal step the global rotational and translational motion is described by the affine parameters. Munderloh et al. [11] propose an image warping method to compensate the global motion of MAVs. This is done using a 2D mesh-based motion compensation technique and improves the result of mosaic creation captured at low altitudes. These methods are designed for images with a wide field-of-view and specific foreground and background. In our approach, a narrow zoom image presenting almost only foreground has to be stabilized, so a correlation-based approach is sufficient.

## 3. Inspection Camera System

We present a mobile distributed system that is capable to present high-resolution images (*detail* images) to an inspector and simultaneously provides low-resolution imagery (*overview* images) with a wide field of view for visual servoing. In contrast to other flying inspection camera systems, we use only a single camera and thus reduce the payload, which in turn increases the flight time. Further, the acquisition time as well as the bandwidth required for transmission are reduced by only exposing and

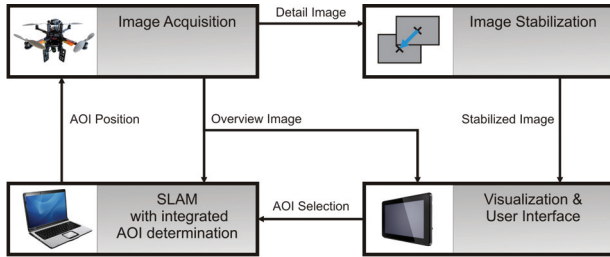


Figure 2. System overview of the inspection camera setup.

transmitting the region of interest for detailed images. Live, full-resolution detail images presented on a tablet computer give the inspector the ability to perform on-line inspection and avoid the need for further post-processing at a later time.

### 3.1. System Overview

For flexible inspection in outdoor environments we use an MAV equipped with a coarsely mechanically stabilized high-resolution camera. The camera is able to change its internal parameters at runtime. Further, the MAV is equipped with an on-board processing unit for image acquisition and parameter management, a wireless 802.11n link for data transmission, and a customer-grade GPS receiver for automatic position-holding.

On the ground we employ a computer for tracking of the object of interest and a mobile tablet computer for visualization and interaction with the inspector. Pose estimation on the MAV is done completely in 3D using a state-of-the-art Simultaneous Localization and Mapping approach named PTAM [9]. This is not only beneficial for tracking the object of interest, but is required anyways for visual servoing.

We want to stress the flexibility and extendability of the proposed system which is gained by the modular design of the framework based on the Robotics Operating System (ROS). Each module such as the 3D pose estimation can be replaced by future, more elaborate algorithms, and the currently manual selection of the Area of Interest (AOI) can be upgraded to an advisor-system in customized applications. Further, the processing pipeline can be easily extended by adding an image processing module to detect faults of the object to inspect automatically.

In the following paragraphs we explain the individual components of the system as presented in Figure 2. The applied theory of projective geometry is based on [6].

### 3.2. Image Acquisition

Using a single camera for overview and detail images could be done by acquiring full-frames and perform down sampling or cropping of the AOI as a preprocessing step of transmission. This kind of acquisition requires a lot of on-board processing power and so results in low frame rates. To get increased performance, we adapt the camera’s hardware settings according to the requested view. Therefore, the camera’s AOI, binning, image mode, and exposure time parameters have to be set while switching from overview to detail image and vice versa. Only the down-sampling for overview images is done prior to the transmission.

### 3.3. Selection of the AOI

The first step is to select the area of interest, in other words the object the inspector wants to inspect. It is specified by a 3D point  $O_i$  that represents the center of the object in the coordinate system  $i$ , as well as by a width  $O_{width}$  and height  $O_{height}$  in pixels. Our approach allows to define the object of interest using the multi-touch capability of the tablet computer. The inspector selects a rectangle around the object of interest by defining the upper left and lower right corner. This rectangle defines  $O_{width}$  and  $O_{height}$  and its center in the current camera frame  $o_c$  the 2D projection of  $O_c$ . To have a complete definition of the object of interest we have to determine the 3D point  $O_w$  in world coordinates. This is done by exploiting the already available 3D localization and mapping data.

Given the sparse 3D map of the environment delivered by Parallel Tracking and Mapping (PTAM) [9], we are able to transform the 3D map points  $X_w$  into the camera’s coordinate system by calculating

$$X_c = H_c^w X_w, \quad (1)$$

where  $X_c$  represents the 3D map points in the camera coordinate system,  $X_w$  the 3D map points in the world coordinate system and  $H_c^w$  the transformation between the camera and world coordinate system as delivered by SLAM. Next, a 3D point  $X_c$  is selected as inlier  $X_{c,inlier}$  if the corresponding 2D point is contained inside the selected object of interest rectangle. This is done by back-projecting  $X_c$  to the camera image plane by

$$x_c = P_c X_c, \quad (2)$$

where  $x_c$  represents a 2D point in sensor coordinates and  $P_c$  is the intrinsic camera matrix. For these in-



Figure 3. For selecting the area of interest (AOI), the inspector selects a rectangular region in the current view on the tablet computer. The depicted selection pipeline is then used to define the object’s center and size in 3D world coordinates.

liers a depth-histogram with  $h_{bins}$  bins is generated based on the z-coordinates of  $X_{c,inlier}$ . Under the assumption that the object of interest is not occluded in the image used for defining it, the first peak in the histogram exceeding a noise removal threshold  $h_{thresh}$  is taken as the coarse depth of interest. The center of the object of interest in camera coordinates  $O_c$  is now defined by the mean of all 3D points of the selected histogram bin and transformed back to world coordinates, so

$$O_w = H_c^c O_c. \quad (3)$$

The entire workflow for selecting the area of interest is depicted in Figure 3.

### 3.4. Tracking of the AOI

For tracking the area of interest we first determine the current camera pose  $H_c^w$  in the same coordinate system where the previously selected object center  $O_w$  is defined. This is achieved by exploiting PTAM [9], which delivers both the visual map and the estimated pose. We transform  $O_w$  to the camera coordinate system, and subsequently back-project it to the camera image plane as

$$o_c = P_c H_c^w O_w. \quad (4)$$

This is also depicted in Figure 4. To save bandwidth, computational efforts, and to reduce jitter, we only update the area of interest on the MAV if its distance to the position in the previous camera coordinate system exceeds  $d_{thresh}$ .

### 3.5. Image Stabilization

The output of the previously described steps is a high-resolution cropped image at the estimated region of interest, based on 3D pose estimation from natural landmarks. While this pose estimate is typically quite accurate, the cropped image is still shifted by a few pixels from frame to frame. Presenting a sequence of such images would lead to an unsteady visualization and requires more attention of the inspector to perform his task. Therefore, the images are

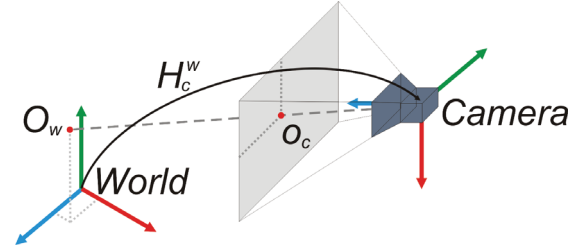


Figure 4. Geometric relationship of the object of interest  $O_w$  and its back-projection  $o_c$  in the current image.

stabilized according to their (stabilized) predecessor. For the correction of the movements we calculate the normalized cross correlation (NCC) and shift the image by the offset of the best match only if the correlation value exceeds a matching threshold  $m_{thresh}$ . Otherwise, the image is placed at its original position because of significant changes in the image. Finally the stabilized image is cropped by a fixed offset  $b$  to remove borders caused by the shift, and this *detail* image is sent to the visualization device.

We explicitly avoid using feature-based approaches in this step. Due to the narrow field of view of the detailed image patches there is an inherent lack of features in homogeneous regions, so feature-based stabilization typically fails. Additionally, detecting larger motions is difficult in the detail images. In contrast, template matching is well suited to correct for the small motions between frames and automatically detects large motions as described above.

## 4. Experiments

We perform several indoor and outdoor experiments using a single industrial camera (IDS UI-1490SE) with a maximum resolution of  $3840 \times 2748$  px and an 8 mm lens. The camera is mounted on a "Pelican" quad-rotor MAV from Ascending Technologies and connected via USB to the on-board computer, an Intel Atom 1.6 GHz processor with 1 GB RAM. The communication to ground is provided by a wireless 802.11n link. A 2.8 GHz quad-core processor serves as ground station. The



Figure 5. Comparison of image acquisition methods: (a) shows a full-resolution, full-frame image with considerable motion blur and rolling shutter effects. In contrast, our approach delivers (b) a binned full-frame overview image used for visual servoing and (c) a cropped full-resolution detail image used for inspection.

user interface is implemented on an Android-based NVIDIA Tegra3 prototype tablet that is connected by a wireless link as well.

In the beginning of the inspection task, an initial map of the environment is generated using PTAM [9]. To achieve correct scaling of the map an ARToolkitPlus marker [13] is placed in the scene and used as reference for setting the map’s origin and scaling. Then, the MAV is (currently manually) navigated to the point of interest for hovering. Using the tablet computer, the inspection object can be selected in the overview image by multi-touch interaction and its 3D position is automatically calculated as described in Section 3.3. After this initialization steps the inspection object is automatically tracked in consecutive frames.

In the following experiments we fixed all parameters to the following values. For the overview image the camera is set to a resolution of  $3840 \times 2720$  px with  $4\times$  binning and  $2\times$  Gaussian pyramid down-sampling, which results in an output image size of  $480 \times 340$  px. The detail image is variable in size depending of the defined region using the full-resolution of the sensor, which means neither binning nor down-sampling is activated. Other values are set to  $h_{thresh} = 0.2$  and  $h_{bins} = 10$  for depth histogram creation,  $d_{thresh} = 10$  mm for AOI updates, and finally  $m_{thresh} = 0.95$  and  $b = 40$  px for stabilization.

#### 4.1. Full-Frame vs. Preview and Zoom Images

In this experiment we show the difference between capturing images at full-resolution and our proposed approach of time-interleaved overview and detail images. Therefore, we measured the frame-rate of the individual approaches under the assumption that the region of interest has a size of  $600 \times 480$  px and is required at 1 fps. For full-frame, non-binned im-

ages we achieve on-board acquisition rates of 2 fps and a transmission rate of 0.67 fps using the wireless link. These low frame rates are neither suitable for tracking on-board nor on the ground. In comparison, given a transmission rate of 1 fps for detail images our approach achieves a maximum transmission rate of 6 fps for overview images.

Besides the low frame rate of full-frame imagery, rolling shutter artifacts are clearly visible (see Figure 5(a)). Exposing a full-frame using a rolling shutter requires more time than only exposing an AOI or a binned frame. Therefore, reasonable artifacts, caused on the one hand by horizontally moving objects and on the other hand by vibrations of the MAV while capturing, are visible in the full-frame imagery. In contrast the images acquired using our approach as shown in Figure 5(b) and 5(c) are suitable for tracking, visual servoing, and inspection, and they are streamed at adequate frame rates.

#### 4.2. Switching Times

For choosing the proper ratio between overview and detail images, it is important to know the timings of the hardware used; especially the time required for switching the parameters on the camera hardware. Therefore, we generate a capturing sequence of three overview images followed by one detail image. The detail image is captured at full-resolution using an image size of  $600 \times 480$  px. We record the timings of the individual steps for 30 capturing sequences to average the timings. We measured an average of 0.17 s for a switch to the detail image settings and 0.19 s to set the overview image parameters. The switching time results from setting AOI, exposure, binning, and image mode. The difference in time arise from an additionally setting of the binning parameter for preview images. A complete illustration of the indi-

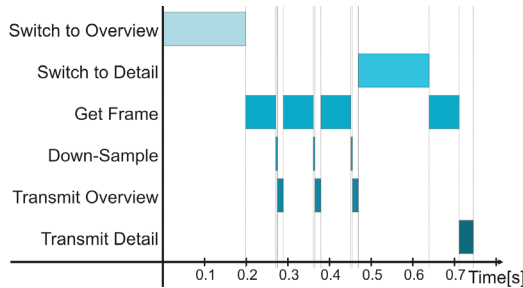


Figure 6. Timings of the individual steps required for a single capturing sequence with three interleaved overview images. It can be seen that switching the parameters takes a considerable amount of time and has to be used rarely.

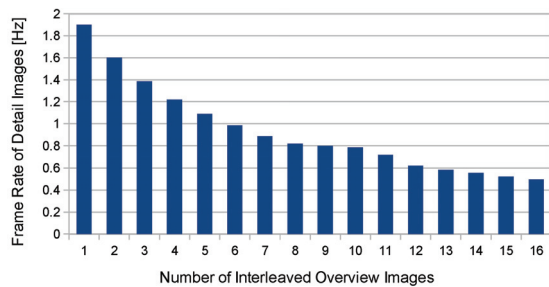


Figure 7. Effect on the measured detail image frame rate given an increasing number of interleaved overviews.

vidual timings is depicted in Figure 6. It can be seen that the switching of the parameters requires a considerable amount of time and has to be used rarely.

### 4.3. Overview Image to Detail Image Ratio

Based on the measured switching times we are able to evaluate the influence of different overview to detail image ratios to maximize the through-put of the wireless link. Therefore, we vary the number of interleaved overview images from 1 to 16 and measure the frame rates of the detail image as shown in Figure 7. Under the constraint of getting an update of the detail image at least every second, which is suitable for an inspection task, at most six overview images can be sent in between. These six overview images are typically sufficient to perform visual servoing.

### 4.4. Impact of Stabilization

Image stabilization is necessary to cope with small pose estimation uncertainties from visual and inertial sensors. We evaluate our stabilization approach by fixing the MAV on the ground and simply switching it on. Then, the coarse mechanical stabilization of the pan-tilt unit tries to stabilize the camera according to



Figure 8. A series of large view point changes. In the top row the overview images are depicted, and in the lower row the tracked detail images. It can be seen that an inspection of one object from different view points is possible without reselecting the object of interest.

the output of the inertial measurement unit, which is noisy. This leads to a slightly shaking image before software stabilization. Our correlation-based stabilization routine is able to deliver a stable image and corrects the positioning errors of the mechanical stabilization unit and the tracking algorithm. The root mean squared (RMS) error of the L2-norm of the distance between stabilized and unstabilized image results in 6.43 px, and the visually evaluated result after stabilization does not show any jitter.

### 4.5. Large Viewpoint Variation

Pose estimation in 3D is beneficial for inspection, as the object of interest can be successfully tracked even in case of large viewpoint changes. This typically happens when the inspector navigates the MAV around an object of interest. In Figure 8, a series of four significant view point changes is depicted, representing the detail images below the according overview images. It can be seen that the detail image shows the same object from different perspectives, although the background changes considerably. This has the advantage that the inspector just has to select the object of interest once, or the object could even be detected automatically for a specific application.

## 5. Conclusion and Future Work

In this work we have presented an airborne inspection system which uses a single camera. We acquire high-resolution detail images at a low frame-rate for inspection and interleave several wide field of view low-resolution overview images suitable for visual servoing. By defining the object of interest in 3D, object tracking exploits the same pose estimation routine as required for visual servoing and thus saves computation time and effort. Further, our method

allows large changes of the view point and enables the inspector to inspect a single object from different perspectives without redefinition of the object of interest. Small accumulated localization errors of the image acquisition pipeline are corrected using a correlation-based image stabilization approach, and large viewpoint changes are automatically detected to re-initialize stabilization.

In future work we would like to evaluate the quantitative influence of interleaved image acquisition on visual navigation and inspection. Moreover, we will customize the inspection system by adding an automatic inspection routine to propose possible application-specific defects. Depending on the inspection task our current support of the inspector could be further improved by adding a point of interest detector in the AOI selection step.

## Acknowledgments

This work has been supported by the Austrian Research Promotion Agency (FFG) project FIT-IT Pegasus (825841/10397). We thank NVIDIA for providing a Tegra3 prototype tablet.

## References

- [1] M. Achtelik, M. Achtelik, S. Weiss, and R. Siegwart. Onboard IMU and monocular vision based control for MAVs in unknown in- and outdoor environments. In *IEEE International Conference on Robotics and Vision (ICRA)*, 2011. 1, 3
- [2] S. Auberger and C. Miro. Digital video stabilization architecture for low cost devices. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, 2005. 3
- [3] A. J. Davison. Real-time simultaneous localisation and mapping with a single camera. In *IEEE International Conference on Computer Vision (ICCV)*, 2003. 3
- [4] A. J. Davison, I. Reid, N. Molton, and O. Stasse. MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 29(6):1052–1067, 2007. 3
- [5] J. Gong, C. Zheng, J. Tian, and D. Wu. An image-sequence compressing algorithm based on homography transformation for unmanned aerial vehicle. In *International Symposium on Intelligence Information Processing and Trusted Computing (IPTC)*, 2010. 3
- [6] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000. 4
- [7] S.-C. Hsu, S.-F. Liang, and C.-T. Lin. A robust digital image stabilization technique based on inverse triangle method and background detection. *IEEE Transactions on Consumer Electronics*, 51(2), 2005. 3
- [8] J. Katrasnik, F. Pernus, and B. Likar. A survey of mobile robots for distribution power line inspection. *IEEE Transactions on Power Delivery*, 25(1), 2010. 2
- [9] G. Klein and D. W. Murray. Parallel tracking and mapping for small AR workspaces. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 225–234, 2007. 3, 4, 5, 6
- [10] H. Meuel, M. Munderloh, and J. Ostermann. Low bit rate roi based video coding for hdtv aerial surveillance video sequences. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011. 2
- [11] M. Munderloh, H. Meuel, and J. Ostermann. Mesh-based global motion compensation for robust mosaicking and detection of moving objects in aerial surveillance. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011. 3
- [12] M. Salous, D. Pycock, and G. Cruickshank. Cbit - context-based image transmission. *IEEE Transactions on Information Technology in Biomedicine*, 5(2):159–170, 2001. 3
- [13] D. Wagner and D. Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In *Proceedings of Computer Vision Winter Workshop (CVWW)*, 2007. 6
- [14] B. Wang, X. Chen, Q. Wang, L. Liu, H. Zhang, and B. Li. Power line inspection with a flying robot. In *1st International Conference on Applied Robotics for the Power Industry (CARPI)*, 2010. 2
- [15] A. Wendel, A. Irschara, and H. Bischof. Natural landmark-based monocular localization for MAVs. In *IEEE International Conference on Robotics and Vision (ICRA)*, 2011. 1, 3
- [16] A. Wendel, M. Maurer, A. Irschara, and H. Bischof. 3D Vision Applications for MAVs: Localization and Reconstruction. In *International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, 2011. 3
- [17] K. W. Weng and M. Abidin. Design and control of a quad-rotor flying robot for aerial surveillance. In *4th Student Conference on Research and Development*, 2006. 2
- [18] J.-M. Zheng, D.-W. Zhou, and J.-L. Geng. Roi progressive image transmission based on wavelet transform and human visual specialties. In *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, volume 1, 2007. 2