JOURNAL OF UNMANNED AERIAL SYSTEMS

TECHNICAL PAPER

STITCHED PANORAMAS FROM LOW-COST AIRBORNE VIDEO CAMERAS

Camille Goudeseune
Computer Systems Analyst, Illinois Simulator Laboratory,
University of Illinois at Urbana-Champaign

Abstract

Effective panoramic photographs are taken from vantage points that are high. High vantage points have become easier to reach as the cost of camera-equipped quadrotor helicopters has dropped to nearly disposable levels. The low-quality video recorded by these cameras can be converted into still panoramas, whose quality and resolution are comparable to images captured from remote-controlled aircraft that are much larger and thus have much greater flight risks.

Introduction

High-quality panoramic photographs (5 to 10 megapixels, with good contrast, sharpness, and gamut) can now be acquired from aircraft under 100 grams. Such aircraft pose much less risk than aircraft carrying higher-quality cameras. An aircraft under 250 g that impacts a bystander has a less than 1% chance of inflicting injury at or beyond severity score 3 of the Abbreviated Injury Scale (Gennarelli & Wodzin, 2008; FAA, 2016). On the other hand, a typical quadrotor carrying a popular airborne camera such as those manufactured by GoPro (75 g to 150 g) has a flying weight of at least 450 g to 900 g, based on a payload fraction of 0.17 to 0.2. (For commercial rotorcraft, this payload fraction holds over three orders of magnitude (Khromov & Rand, 2006, fig. 10), or over five orders of magnitude when considering DARPA's goal for UASs under 10 g (Hylton, Martin, et al., 2012).) Such "category 2" aircraft need safety constraints not needed by lighter aircraft: formal tests to verify that, in likely failure modes, impact energy is less than a set threshold; an analysis of rotor impact; and an operating manual that includes requirements for flight near bystanders, at least 6 m above or 3 m laterally away from them (FAA, 2016). The categories for even heavier aircraft have correspondingly stricter constraints, such as operator qualifications, restrictions on flying over crowds, and creating and following a risk mitigation plan. In short, aerial photography can now present little risk to both bystanders and the UAS itself, which can be inexpensive enough to replace outright if it is lost or damaged.

Capturing video from sub-100 g quadrotors is commonplace (Chen, 2015), but no reports have been published about capturing still images. This paper offers a complete set of techniques for acquiring high-quality panoramas from such aircraft. In order, the sections of this paper describe how to extract still frames from these videos; maneuver effectively to avoid motion parallax; suppress artifacts due to poor camera quality; cope with strong winds and motion blur; and record simultaneously from multiple cameras (on multiple quadrotors) to broaden a panorama. These techniques are all suitably simple and inexpensive.

Although many aerial tasks rely on cameras, other payloads are not unknown. Insofar as the techniques presented here generalize beyond cameras, they may inspire others to try lighter, simpler sensors on lighter, simpler aircraft. Taken together, all of this reduces the risks inherent to flight, widens a task's range of deployment, and makes aerial sensing more accessible to grassroots movements, nonprofits, and organizations in developing countries.

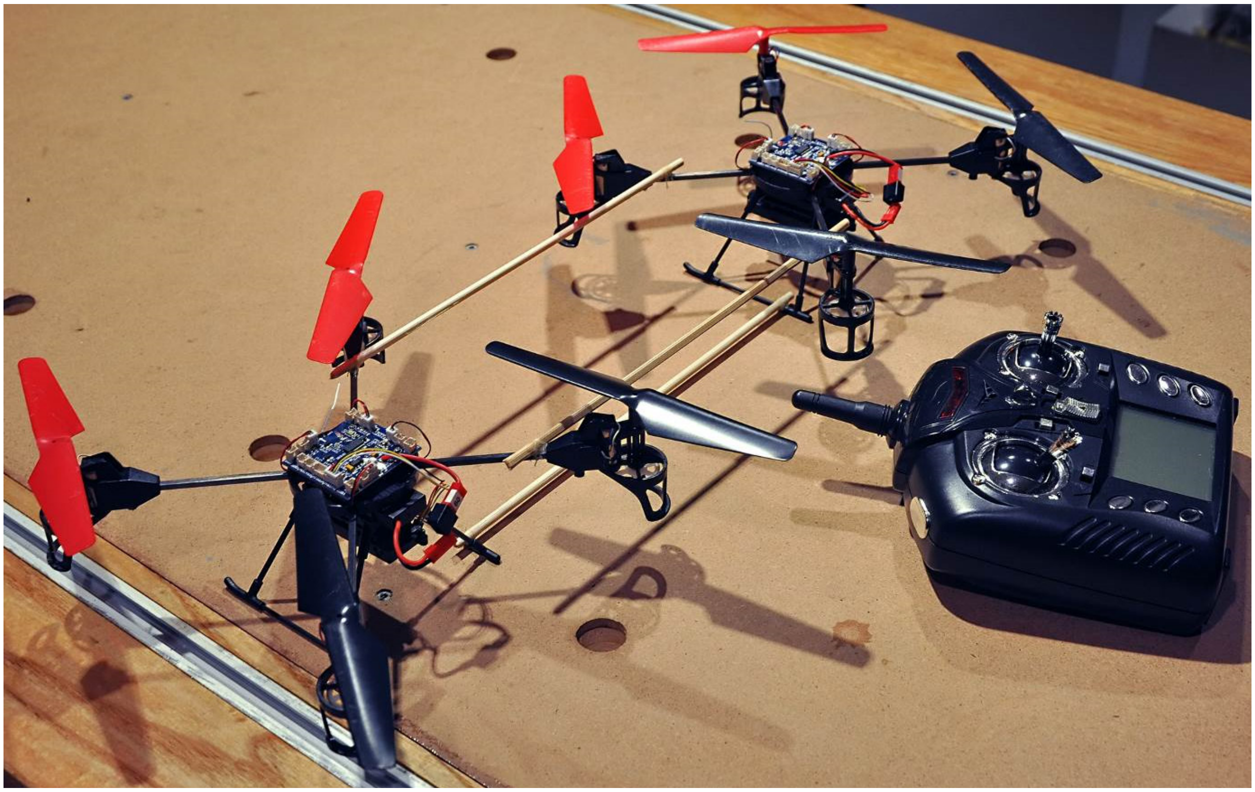

Figure 1. Two videocamera-equipped quadrotors, with a shared radio-control transmitter.

Quadrotors

During the 1990's, electric power for radio-controlled aircraft improved to match the power of piston engines. At the same time, electric drivetrains kept their advantages of reliability, low vibration, and mechanical simplicity. Electric power became mainstream. Over the next decade, the enduring consumer preference for mechanical simplicity then led to the quadrotor helicopter, with only four moving parts on the entire aircraft. Pushing almost all of the aircraft's complexity into software made it inexpensive, maintenance-free, and crash-tolerant. On e-commerce web sites, the price of sub-100 g camera-equipped quadrotors, such as those in fig. 1, has fallen below USD 15.

Quadrotors in the range of 100 g are inconspicuous and quiet. Although quadrotors as light as 12 g have become widely available in the past two years, reviews in hobby magazines agree that they are overwhelmed by winds stronger than a few knots. Thus, a flying weight near 100 g may remain optimal for some time. Besides offering stealth, small size also lets the aircraft be carried around more often, to capture images at unplanned opportune moments (fig. 2).

Figure 2. Full 360 degree panorama. Trenton, Ontario, 2013-08-19.

Converting a Video to a Panoramic Image

Panoramas can be extracted from many kinds of video recordings, of course, but particular issues apply to videos recorded by cameras on small quadrotors.

In nonaeronautical contexts, the camera often found on small quadrotors is called a keychain camera or an "808" (Lohr, 2016). Its specifications change monthly, but are roughly: mass 8 g, pixel resolution 640x480 to 1280x800, microSD card storage, 30 or 60 frames per second, fixed focus, and depth of field 10 cm–infinity. Its 2 mm diameter lens performs poorly in low light, so flying at dusk or indoors should be avoided. The lens's narrow field of view means that the pixel resolution matters less than one might think. For example, fig. 2 includes plenty of detail, despite being stitched from a video whose resolution is only 640x480. The final panorama has eight times more pixels than any individual frame from that video.

Keychain cameras save a video file in motion-JPEG format, which is just a sequence of individual JPEG images that happens to be synchronized to a soundtrack (Library of Congress, 2012). Because this format does not exploit inter-frame redundancy, it produces files 3 to 10 times larger than those made with a modern codec. This large size is tolerable, though, because it does not constrain recording—videos of a dozen 5-minute flights easily fit on a modest 8 GB card. In fact, a camera with stronger file compression would drain the aircraft's battery faster, paradoxically decreasing the duration of both a flight and its recording.

Individual frames from the video file can be extracted with software ranging from idiotproof smartphone apps to elaborate Swiss Army knives like FFmpeg (2016). For the latter, a typical command is `ffmpeg -i infile.mov -vcodec png -f image2 %04d.png`. This produces image files named `0001.png, 0002.png`, …, . (Alternatively, if image quality needs no further improvement , the original JPEG frames can be instantaneously extracted: `ffmpeg -i infile.mov -vcodec copy -f image2 %04d.jpg.)`

The resulting collection of video frames usually needs some preparation before stitching. Dropped or missing frames occur with some camera-card combinations, or when the camera's CPU is momentarily too slow for even the simplistic motion-JPEG format. Naive extraction of frames "reconstructs" these missing frames by repeatedly duplicating the previous frame, but such duplication slows down image stitching. Many of these consecutive duplicate frames can be removed by specifying the option `-vsync 0` to FFmpeg. Removing all duplicate frames requires a duplicate-file finder, of which dozens are available for free online. Because these finders use file size as a quick first test for duplication, they are much slower with formats such as `.bmp` and `.ppm` that give every frame the same file size. The `.png` format does not suffer from this, and is thus preferred.

These image files may be further improved as needed by applying the tools discussed later in this paper. When they are ready, they are sent to an automatic image stitcher, such as the free programs AutoStitch (Brown, 2015; Brown & Lowe, 2007) and Image Composite Editor (Microsoft, 2016). The stitcher then produces a single panoramic image (figs. 2, 3, and 8).

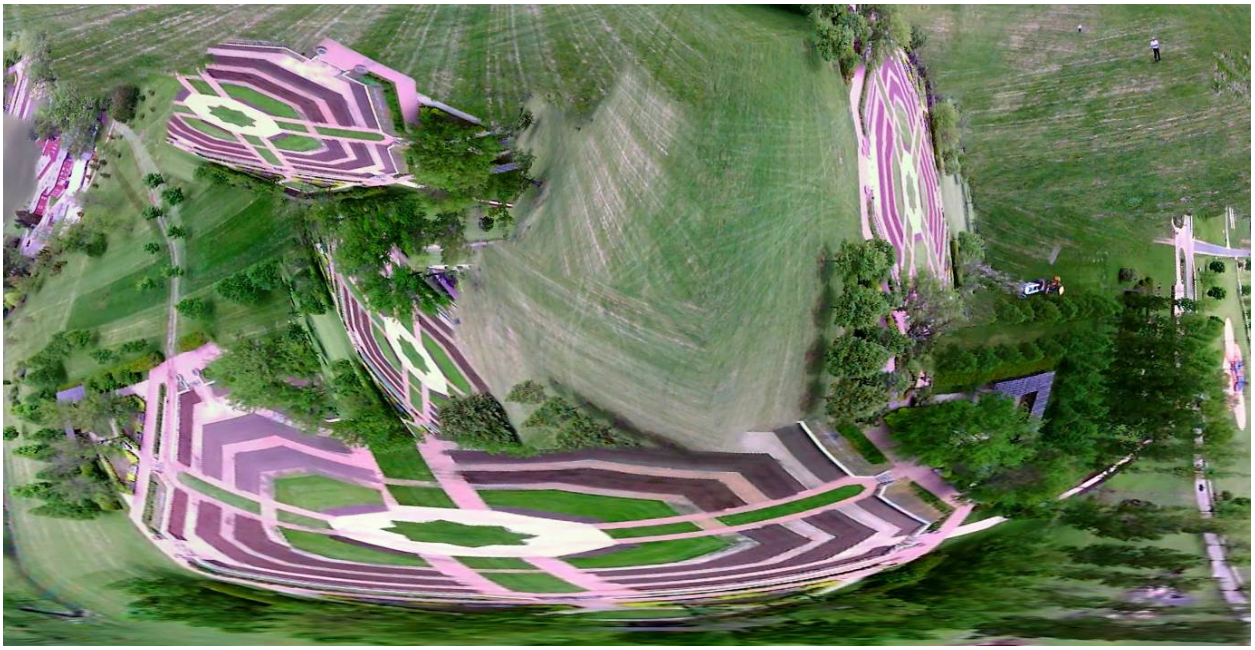

Figure 3. Mis-stitching due to camera movement. UIUC Arboretum, Urbana, Illinois, 2013-05-16.

Flight Paths Optimized for Stitching Images

Stitching software assumes that the images it is given were captured from a single viewpoint. Because a lightweight quadrotor is hardly a stationary tripod, this assumption is spectacularly violated by stitching the video recording of an entire flight (fig. 3). For a coherent panorama, only a subinterval should be stitched.

A convenient way to capture a stitchable subinterval is to pirouette the quadrotor, yawing about a vertical axis while avoiding other rotations or movements (in cinematography, this is the very definition of a pan shot, the abbreviation for panorama). Some drifting is tolerable if the subject is more than about 20 m away, and if the pirouette is less than a full circle. Stitching gains accuracy when frames have more overlap, which happens with slower yaw. The slowest practical yaw for a 100 g quadrotor is about 0.4 rad/s, or 16 s for a full pirouette. Even slower stationary yaws would require more mass or a GPS-enabled autopilot to resist wind gusts. Because 0.4 rad/s is still fast enough to avoid stitching artifacts such as ghosting (Chen & Huang, 2012b; Uyttendaele, Eden, et al., 2001), it is preferred.

Choosing a Stitchable Subinterval

After landing, the video is viewed on a computer to find a brief interval that shows the desired subject. To maximize the panorama's coverage, one stretches the interval's endpoints as far as one can, while restricting it to a single pan (without reversing direction) from a single viewpoint. Although several back-and-forth pans would cover the subject more thoroughly, in practice each pan would be from a slightly different viewpoint, introducing seams like the one in fig. 8 just left of the red truck. The more these viewpoints differ, the more obvious the seams are (fig. 3).

Video Downlinks

After reviewing the videos from a few flights, most pilots develop an intuition for what the quadrotor's camera is seeing. But if the quadrotor's height exceeds that used to capture fig. 2, about 30 m, it becomes almost too small to orient in flight, to aim its camera. (Its body is less visible than a baseball at home plate as seen from first base.) If aiming is a concern, a live video downlink can be added to the quadrotor. But this extra equipment, and lightening of the stock parts to compensate for this payload, costs ten to twenty times more than the original aircraft (Gaffoor, 2013). Also, the many single points of failure of such first-person view (FPV) flight add risk, counter to the whole point of using this kind of aircraft. Even for a 100 g aircraft, the prudent FPV pilot flies nearby enough to revert to line-of-sight control, and maintains situational awareness with a spotter assistant. When flying much more than 30 m away, with or without FPV, it becomes safer to instead fly an aircraft that is large enough to be visible at that distance.

Suppressing Camera Artifacts

A keychain camera's image quality may be poor in several ways: varying brightness, rolling shutter, moire bands, and compression blockiness. Fortunately, these artifacts can be suppressed or eliminated.

Reducing Variations in Brightness and Color

Brightness and color variations in a video are due to the camera's automatic exposure compensation and automatic white balance (Uyttendaele, Eden, et al., 2001). When the view changes suddenly from bright cumulus to shaded terrain, the camera takes a few seconds to correct its exposure. Similarly, when a view of only grass suddenly tilts up to include some sky, a full second elapses before the white-balanced grayish grass becomes bright green again. Frames from such transitions may not be usable for stitching. Reducing such variations requires slower aircraft rotation. After flight it may be too late to correct the transitional frames if color is out of gamut, or if shadows or highlights are clipped (lost detail, in pure black shadows or pure white highlights).

Minimizing Artifacts due to Rolling Shutter

Small cameras often use a "rolling" shutter, which captures an image one scanline at a time, instead of all at once. In other words, different parts of the image correspond to different instants in time. Therefore, moving the camera relative to the subject produces visible warp and skew. As with varying brightness, the first cure is slower aircraft rotation. Also, balancing the propellers with flecks of adhesive tape reduces the mechanical vibration that causes "jello" in video (Graham, 2013).

Unlike varying brightness, though, rolling shutter can be suppressed after flight (Baker, Bennett, et al., 2010; Grundmann, Kwatra, et al., 2012). Rolling shutter repair is included in commercial video software such as Adobe Premiere Pro and Adobe After Effects, and in free video software such as the Deshaker plug-in for VirtualDub (Thalin, 2014; Lee, 2013). However, these tools specialize in inter-frame smoothness, which is not needed for panorama stitching. Worse, they may crop the image (which shrinks the panorama's coverage) or add a black border (which confuses the stitcher). If the border's color can be made transparent, however, commercial stitchers such as Adobe's Photomerge may succeed. Better yet, Deshaker can fill the border with pixels from previous or successive frames, or, when those are unavailable, with colors extrapolated from the current frame.

Avoiding Moire Artifacts

The artifact called a moire pattern consists of undesired bands of hue or brightness (fig. 4), seen in a subject with repetitive detail that exceeds the camera's resolution. If the stitcher tries to match such stripes or bands, which shift from frame to frame as the camera moves slightly, stitching quality deteriorates. This is particularly so for stitchers that match image features by hue as well as by brightness, because a camera sensor's Bayer filter mosiac produces strong hue bands. The pattern is due to foldover at the camera's Nyquist frequency. Better cameras suppress this with anti-alias filters. To avoid moire patterns, then, one must either fly quite far from such a subject, or fly so close that each stripe is at least two pixels wide (for a keychain camera, at most a few hundred stripes visible at once).

Figure 4. Different magenta-cyan moire patterns on three identically corrugated roofs. The roofs differ only in their distance from the camera. UIUC Dairy Cattle Research Unit, 2013-08-01.

Removing Blockiness due to JPEG Compression

Some JPEG frames may be compressed so strongly that a grid appears at the boundary between 8x8 pixel blocks. As with moire patterns, this noise varies from frame to frame, distracting the stitcher from matching common elements across frames. Such grids are also distracting in the final panorama. This artifact is suppressed by the UnBlock algorithm (Costella, 2006; Goudeseune, 2014), which smooths over the boundaries between blocks, but only aggressively enough to reach the same distribution of discrepancies across the block boundaries as is found in the block interiors (fig. 5). The algorithm's tuning-free design prevents it from introducing other visual artifacts.

Tethering a Quadrotor to a Kite

In winds too strong for a lightweight quadrotor, it can nevertheless be given a high vantage point by hanging it from a toy delta-wing kite (span 1.3 m, cost USD 5). Even with the aircraft's four booms removed to prevent the rotors from getting fouled in the kite's tether, the aircraft still operates as a power source and remote control for the camera (fig. 6).

An elaborate Picavet camera suspension (Picavet, 1912; Beutnagel, Bieck, et al., 1995) is inappropriate: building and testing one takes several hours, while a commercial unit costs many times more than the entire quadrotor. On the other hand, just dangling the camera from the kite's tether shakes the camera so much that fewer than one frame in a hundred is stitchable (fig. 7). Happily, the shaking can be dampened by hanging the camera from not one but two points on the tether, at the bottom of a "V." Then one frame in ten has acceptably low motion blur. But this then poses the problem of how to find these still rare frames.

Culling Motion-Blurred Images by Automatic Sorting

Manual culling of frames blurred by camera motion is impractical. To automate this, one can measure how blurred each frame is, and then sort the frames by blurriness. A frame's blurriness is measured simply and thus robustly by re-saving it in JPEG format, with and without first applying a Gaussian blur. The smaller the ratio of the sizes of the two resulting JPEG files, the less effective the Gaussian blur was, and thus the blurrier the original frame was. Because calculating a blur is slow, downsampling beforehand by a factor of four or so greatly speeds up the processing of thousands of frames. (Downsampling also attenuates sharp pixel-block boundaries such as those in fig. 7. This is desirable because those sharp boundaries reduce how well the Gaussian blur approximates the original motion blur—they hide the smooth motion blur behind artificial crisp edges.) One can then build a list of pairs of frame-filenames and blurriness values. After sorting this list by increasing blurriness, the start of the list then yields a set of least blurry frames. A more elaborate deblurring method, culling any video frame that has few sharp edges compared to its neighbors (Chen and Huang, 2012a), fails in the presence of duplicate frames, which we have already discussed.

Of course, a Gaussian blur only approximates a motion blur. But the exact motion blur is a combination of axial rotation and panning, which is too expensive to measure for this quick first pass that culls almost all of the frames. Later passes can use advanced algorithms (Li, Kang, et al., 2010; Cho, Wang, et al., 2012), which not only detect but even remove mild blur by estimating camera motion from consecutive frames—although these again fail for duplicate frames. Such advanced deblurring can also improve videos taken without the help of a kite.

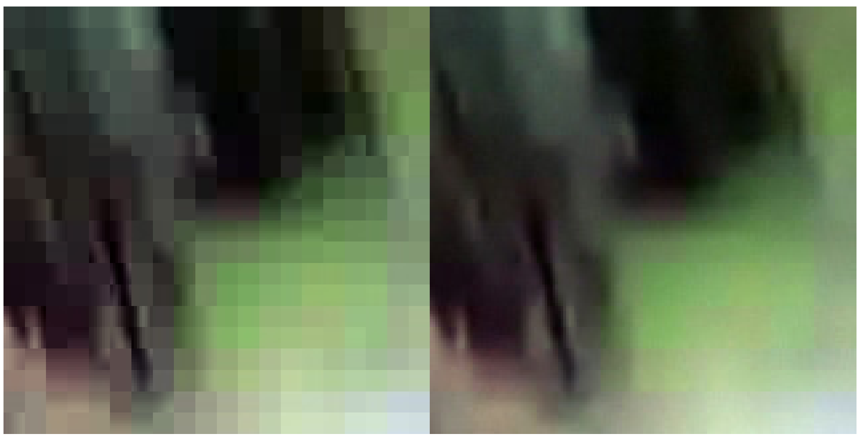

Figure 5. Detail (160x160 pixels) from top right of fig. 7. Left: original. Right: processed by the UnBlock algorithm.

Figure 6. Kite hoisting a rotorless quadrotor-camera (a “nullicopter”), while capturing fig. 7.

Figure 7. Strong motion blur from a kite-suspended camera. Evergreens 5 m to 15 m tall, Okanogan-Wenatchee National Forest, 2013-05-22.

Multiple Cameras

Terrain hillier than what is shown in fig. 2 may need more vertical coverage than what one camera can provide. The same holds for panoramas that need coverage of nearby foreground as well as distant background (fig. 8). In this case, a quadrotor may have enough thrust to carry more than one camera, with each camera pointing in a slightly different direction. This has been proposed for Parrot's AR.Drone quadrotor (400 g, USD 300) (Chen and Huang, 2012a), but no implementations to date have used aircraft under 100 g. More typical is DARPA's ARGUS-IS cluster of several hundred cameras (Vaidya, 2011).

Figure 8. Top: panorama stitched from one camera's frames. Bottom: second camera's frames added. UIUC Large Animal Clinic, 2013-10-09.

If the quadrotor's maneuverability suffers with the extra payload of more cameras, another novel solution is to laterally combine two or more (figs. 1 and 9). (Buying multiple aircraft is an inexpensive way to get spare parts.) The transmitter is unaware that it is controlling more than one quadrotor. The composite aircraft is less maneuverable because the stabilizers in each quadrotor fight each other, and because roll authority is reduced. But the more important controls—pitch, yaw, and thrust—have no reduced authority. As with multiple cameras on one quadrotor, each camera points at a different angle.

Figure 9. Two-camera octorotor, just before capturing fig. 8.

Conclusions and Extensions

High-quality panoramic photos can be captured with a videocamera-carrying quadrotor of startlingly small size and cost, thanks to multiple stages of software post-processing. These stages can be applied to whichever aspects of a particular panorama need improving.

Multiple cameras can record stereoscopic video, especially when widely spaced on an octorotor. The audio recorded with each camera's rudimentary microphone suffices to synchronize the individual monoscopic video recordings.

A stereoscopic panorama can be made with only one camera, by recording two partial pirouettes from nearby locations. A typical sequence of maneuvers would be a half pirouette, forward flight for a few seconds, and then a half pirouette in the opposite direction. Each half pirouette is stitched into its own monoscopic panorama; these two are then combined.

Although the data recorded by lightweight sensors carried by lightweight aircraft may be inferior to what is recorded by standard equipment, post-flight software can ameliorate this. Furthermore, in some cases the reduced cost and reduced flight risks of lightweight equipment may make it the only practical way to collect data at all.

References

Baker, S., Bennett, E., Kang, S. B., & Szeliski, R. (2010). "Removing rolling shutter wobble," in Proc. IEEE Computer Vision and Pattern Recognition, pp. 2392-2399. http://dx.doi.org/10.1109/CVPR.2010.5539932

Brown, M. (2015). AutoStitch. www.cs.bath.ac.uk/brown/autostitch/autostitch.html.

Brown, M., & Lowe, D. (2007). "Automatic panoramic image stitching using invariant features," in Intl. J. Computer Vision 74(1), 59-73.

Chen, J. (2015). The Micro RTF Quadcopters Thread. www.rcgroups.com/forums/showthread.php?t=1701910.

Chen, J., & Huang, C. (2012a). "Ghosting elimination with A* seam optimization in image stitching," in Proc. Intl. Conf. on Information Security and Intelligence Control, pp. 214-217.

Chen, J., & Huang, C. (2012b). "Image stitching on the unmanned air vehicle in the indoor environment," in Proc. Soc. Instrument and Control Engineers, pp. 402-406.

Cho, S., Wang, J, & Lee, S. (2012). "Video deblurring for hand-held cameras using patch-based synthesis," in ACM Trans. Graph. 31(4). http://dx.doi.org/10.1145/2185520.2185560.

Costella, J. (2006). The UnBlock algorithm. http://johncostella.webs.com/unblock/unblock_paper.pdf.

FAA. (2016). Micro unmanned aircraft systems aviation rulemaking committee final report. www.faa.gov/uas/publications/media/Micro-UAS-ARC-FINAL-Report.pdf.

FFmpeg. (2016). FFmpeg. www.ffmpeg.org.

Gaffoor, F. (2013). Micro FPV HD quad: WLtoys v929/949/959 FPV HD. www.rcgroups.com/forums/showthread.php?t=1809332.

Gennarelli, T., & Wodzin, E., eds. (2008). The abbreviated injury scale 2005 - update 2008. Assn. for the Advancement of Automotive Medicine.

Goudeseune, C. (2014). UnBlock. https://github.com/camilleg/unblock.

Graham, J. (2013). "Multirotors for the beginner, part two," in Model Aviation 39(5), 77-79.

Grundmann, M., Kwatra, V., Castro, D., & Essa, I. (2012). "Calibration-free rolling shutter removal," in Proc. IEEE Conf. Computational Photography, pp. 1-8. http://dx.doi.org/10.1109/ICCPhot.2012.6215213.

Hylton, T., Martin, C., Tun, R., & Castelli, V. (2012). "The DARPA nano air vehicle program," in Proc. AIAA Aerospace Sci. Meeting.

Khromov, V., & Rand, O. (2006). "Design trends for rotary-wing unmanned air vehicles," in Proc. Intl. Congress of the Aeronautical Sciences.

Lee, A. (2013). VirtualDub. http://virtualdub.org.

Li, Y., Kang, S., Joshi, N., Seitz, S., & Huttenlocher, D. (2010). "Generating sharp panoramas from motion-blurred videos," in Proc. IEEE Computer Vision and Pattern Recognition, pp. 2424-2431.

Library of Congress. (2012). Motion JPEG 2000 File Format. www.digitalpreservation.gov/formats/fdd/fdd000127.shtml.

Lohr, C. (2016). 808 Car keys micro camera, micro video recorder, review. www.chucklohr.com/808.

Microsoft Research. (2016). Image composite editor. http://research.microsoft.com/ivm/ice.

Thalin, G. (2014). Deshaker. www.guthspot.se/video/deshaker.htm.

Uyttendaele, M., Eden, A., & Szeliski, R. (2001). "Eliminating ghosting and exposure artifacts in image mosaics," in Proc. IEEE Computer Vision and Pattern Recognition, pp. II/509-II/516.

Vaidya, S. (2011). "From video to knowledge," in Science and Technology Review, pp. 8-11. https://str.llnl.gov/Apr-May11/pdfs/4.11.1.pdf.