NRC Research Press

ARTICLE

Airborne seismic data acquisition using stereo vision

Thomas Rapstine, Paul Sava, and Elias Arias

Abstract: Environmental impact and the high cost of acquiring land seismic data are major factors to consider when designing a seismic survey. We explore means to quickly record seismic data without disturbing, or contacting, the ground surface, while reducing environmental impact and cost. Recent developments in computer vision techniques and unmanned aerial vehicle (UAV) technology lead us to propose passively observing ground displacement with a UAV-borne stereo video camera system. The recovered displacement is represented as a time-varying probability density function (PDF) of ground displacement. Using this PDF, we can estimate the uncertainty of the displacement measurements. We conclude that currently available camera and UAVs may be used to measure sub-millimetre ground displacements, with associated uncertainties.

Key words: seismic acquisition, airborne geophysics, computer vision.

Résumé : L'impact sur l'environnement et le coût élevé de l'acquisition de données sismiques sur le terrain sont les principaux facteurs à considérer au moment de concevoir un levé sismique. Nous explorons des moyens pour pouvoir rapidement enregistrer des données sismiques sans perturber ni entrer en contact avec la surface du sol, tout en réduisant l'impact sur l'environnement et le coût. Les récents développements au niveau des techniques de vision par ordinateur et de technologie des véhicules aériens sans pilote (UAV) nous poussent à proposer d'observer de façon passive le déplacement du sol à l'aide d'un système de caméra vidéo stéréo à bord d'un UAV. Le déplacement récupéré est représenté par une fonction de distribution de probabilité du déplacement du sol variant en fonction du temps. En utilisant cette fonction de distribution de probabilité, nous pouvons évaluer l'incertitude des mesures de déplacement. Nous concluons que les appareils photographiques et les UAV actuellement disponibles peuvent être utilisés pour mesurer des déplacements du sol sous-millimétriques, ainsi que les incertitudes qui y sont liées. [Traduit par la Rédaction]

Mots-clés : cquisition sismique, levés géophysiques aériens, vision par ordinateur.

Introduction

Geophysical survey design requires one to consider the environment, access, and acquisition cost. When operating in environmentally sensitive areas, it may not be possible to conduct a ground-based survey without disturbing an ecosystem. In addition, in remote locations a survey may not be feasible because of access or other logistical concerns.

Received 2 August 2017. Accepted 2 November 2017.

T. Rapstine and P. Sava. Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA.
E. Arias. Spectraseis Inc., 1600 Stout St., Denver, CO 80202, USA.
Corresponding author: Thomas Rapstine (e-mail: trapstin@mines.edu).
Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from RightsLink.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Financial limitations also influence data acquisition feasibility. Cost may be reduced by optimizing survey design parameters, such as seismic source geometry, source signal length and frequency content, and receiver spacing. Technological advancements can also reduce costs and potentially lead to new opportunities for exploration; an excellent example is distributed acoustic systems (Bostick III 2000; Mestayer et al. 2011; Mateeva et al. 2012; Daley et al. 2013; Mateeva et al. 2013, 2014). We advocate, in this paper, that recent developments in computer vision and robotics also have impact potential for seismic data acquisition, leading towards fast, low cost, and environmentally friendly airborne seismic surveying.

Monitoring motion without contacting a surface can be performed using passive or active systems. Laser Doppler vibrometers are examples of active systems, which use known source signals and the phase of the return signals to deduce the distance the signal travelled to a target. Laser Doppler vibrometers have been used in the past to investigate the feasibility of remotely detecting ground motion from seismic waves (Berni 1991, 1994). Other active systems use microwaves, and interferometry, to deduce vibrations and resonant frequencies in a structural engineering context, such as in Stanbridge et al. (2000). In contrast to active systems, passive systems do not require a source for deducing motion and leverage high speed video cameras and ambient lighting. Here we focus on passive methods for deducing motion.

We advocate measuring ground displacement as a function of time using stereo vision theory. As our left and right eyes allow for depth perception, two images taken from laterally offset cameras allow for a distance estimate. Stereo vision has been used in the past to passively deduce distance using two images taken from laterally offset cameras (Quam et al. 1972). We measure the lateral shift, or disparity, between points observed in the two images. As we show later, this disparity is inversely proportional to the distance from the camera to the points observed in the images. Therefore, the process of finding ground displacement from stereo images requires us to accurately compute perceived shifts between two images. The stereo vision process is discussed in more detail in the section entitled "Theory".

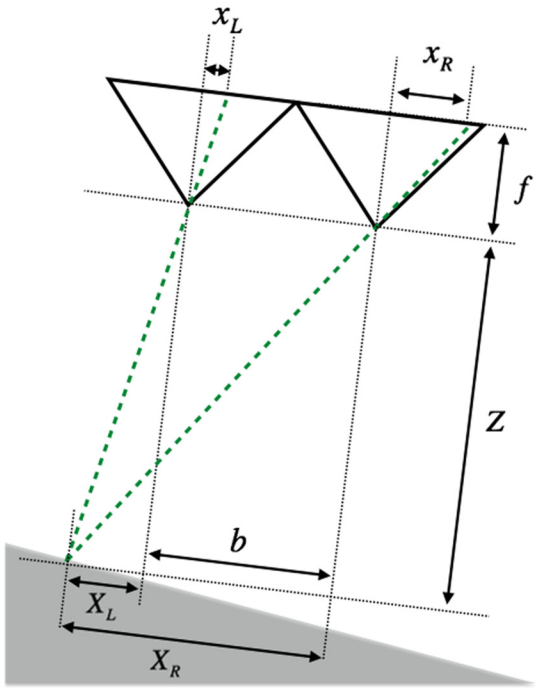
In this paper, we demonstrate and detail a method for sampling a seismic wavefield, represented by ground displacements, that has potential to reduce costs while providing new opportunities for geophysical exploration or earthquake monitoring. First, we clarify the stereo vision theory used to determine ground displacement variations with time. Using the collection of ground displacement values, we construct a probability density function (PDF) of ground displacement, from which we deduce the uncertainty of our measurement. Lastly, we conduct a realistic computer graphics simulation of a real-life earthquake signal, demonstrating the feasibility of using stereo videos to recover sub-millimetre displacement signals, with associated uncertainties, from a moving airborne platform.

Theory

We aim to passively measure ground displacement, without touching the ground, and obtain an uncertainty estimate of our measurement. In this section, we provide a brief summary of how stereo vision is used to acquire a ground displacement measurement with associated uncertainty. The stereo vision theory we summarize is common in standard computer vision literature (Szeliski 2011). Stereo videos are represented as a sequence of frames taken from two offset cameras (left and right), as shown in Fig. 1. This geometry allows one to acquire the distance to points viewed by the stereo cameras using geometric relations based on similar triangles and triangulation.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Fig. 1. Schematic of stereo cameras viewing the ground from above. Dashed paths show where a point on the ground projects on the left and right cameras, each with focal length f. The cameras optical axes are aligned and separated by b.



We begin by defining the origin of three coordinate frames representing the origins of the world, left camera, and right camera. It is our goal to consistently represent a PDF of ground position in the world coordinate frame using images from stereo cameras. In the world coordinate frame, we denote the world and left camera origins as ${}^{\mathrm{w}}\boldsymbol{o}_{\mathrm{w}} = [0, 0, 0]^{\mathrm{T}}$ and ${}^{\mathrm{w}}\boldsymbol{o}_{\mathrm{l}} = [0, 0, -h]^{\mathrm{T}}$, respectively. We adopt the notation where preceding superscripts denote the coordinate frame of a point and subscripts denote the label of a specific point as w, l, and r for the world, left camera, and right camera coordinates, respectively. The z-axis of the world coordinate frame points down, and a left camera is placed at a height h above the world origin. The right camera origin is placed a distance b away from the left camera in the x-direction and is noted by ${}^{\mathrm{l}}\boldsymbol{o}_{\mathrm{r}} = [b, 0, 0]^{\mathrm{T}}$. We represent transformations between an arbitrary coordinate system a to another coordinate system b using a 4×4 matrix, ${}^{b}_{a}\mathbf{H}$. This matrix may be obtained by augmenting the rotation matrix from coordinate system a to b, denoted ${}^{b}_{a}\mathbf{R}$, with the origin of a in b, denoted ${}^{b}\mathbf{t}_{a,\mathrm{org}}$

$$(1) \quad {}^{b}_{a}\mathbf{H} = \begin{bmatrix} {}^{b}_{a}\mathbf{R} & {}^{b}\mathbf{t}_{a,\mathrm{org}} \\ \mathbf{0} & 1 \end{bmatrix}.$$

We use a homogeneous point representation to apply rotation and translation of a point using a single matrix–vector multiplication. A point $\boldsymbol{x}$ in 3D space may be represented in homogeneous coordinates as a four-element point $\widetilde{\boldsymbol{x}}$ by appending a fourth arbitrary element to $\boldsymbol{x}$. We represent a point $\boldsymbol{x}$ in homogeneous coordinates using the notation $\widetilde{\boldsymbol{x}}$, and note that homogeneous coordinates that are scalar versions of one another are

considered equivalent points in 3D space. We may obtain the original coordinate from the homogeneous coordinate by dividing by the fourth element:

$$(2)\quad \tilde{\boldsymbol{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} \frac{x_1}{x_4} \\ \frac{x_2}{x_4} \\ \frac{x_3}{x_4} \end{bmatrix}.$$

With homogeneous coordinates, we can transform the right camera origin from the left camera frame to the world frame by

$$(3)\quad {}^{\mathrm{w}}\tilde{\boldsymbol{o}}_{\mathrm{r}} = {}^{\mathrm{w}}_{\mathrm{l}}\mathbf{H}\ {}^{\mathrm{l}}\tilde{\boldsymbol{o}}_{\mathrm{r}}.$$

In general, a point in the right camera frame is transformed into the world coordinate frame using

$$(4)\quad {}^{\mathrm{w}}\tilde{\boldsymbol{x}} = {}^{\mathrm{w}}_{\mathrm{r}}\mathbf{H}\ {}^{\mathrm{r}}\tilde{\boldsymbol{x}}.$$

Stereo vision allows us to determine the 3D location of a point $\boldsymbol{p}$ observed using the left and right cameras. We can represent the point in the left or right coordinate frames as ${}^{\mathrm{l}}\boldsymbol{p} = [X_{\mathrm{l}}, Y_{\mathrm{l}}, Z_{\mathrm{l}}]^{\mathrm{T}}$ or ${}^{\mathrm{r}}\boldsymbol{p} = [X_{\mathrm{r}}, Y_{\mathrm{r}}, Z_{\mathrm{r}}]^{\mathrm{T}}$, respectively. We note that $Z_{\mathrm{l}} = Z_{\mathrm{r}} = Z$ because the camera origins are only separated in the x-direction of the right camera frame. Using stereo vision, we find ${}^{\mathrm{r}}\tilde{\boldsymbol{p}}$ then transform the point to the world coordinate frame using ${}^{\mathrm{w}}\tilde{\boldsymbol{p}} = {}^{\mathrm{w}}_{\mathrm{r}}\mathbf{H}\ {}^{\mathrm{r}}\tilde{\boldsymbol{p}}$. The 3D location of the point ${}^{\mathrm{r}}\boldsymbol{p}$ is found using the pixel coordinates where the point is observed in the left and right cameras, denoted $(x_{\mathrm{l}}, y_{\mathrm{l}})$ and $(x_{\mathrm{r}}, y_{\mathrm{r}})$, respectively. Projecting a 3D point, such as ${}^{\mathrm{r}}\boldsymbol{p}$, onto a 2D image can be described using a pinhole camera model, which assumes that all 3D points project onto the image plane of a camera through a single point, or pinhole. The model is appropriate to use when a camera has been calibrated to remove lens distortion, which is commonly performed on images (we assume all images are calibrated). Figure 1 shows a point projecting onto the left and right image planes using the pinhole camera model. We use similar triangle geometry to express the left and right x-pixel coordinates of the point as a function of the camera focal length f, in pixels, and the distance to the point Z

$$(5)\quad \frac{x_{\mathrm{l}}}{f} = \frac{X_{\mathrm{l}}}{Z} \rightarrow x_{\mathrm{l}} = \frac{fX_{\mathrm{l}}}{Z} \qquad \frac{x_{\mathrm{r}}}{f} = \frac{X_{\mathrm{r}}}{Z} \rightarrow x_{\mathrm{r}} = \frac{fX_{\mathrm{r}}}{Z}.$$

The shift of the 3D point, as observed by left and right cameras, known as disparity, is then

$$(6)\quad d = x_{\mathrm{r}} - x_{\mathrm{l}} = \frac{fX_{\mathrm{r}} - f(X_{\mathrm{r}} - b)}{Z} = \frac{fb}{Z}.$$

We note that the shift is an integer because the pixel coordinates are integers. The distance, in the z-direction extending from the cameras, can be used to deduce the x- and y-coordinates of ${}^{\mathrm{r}}\tilde{\boldsymbol{p}}$

$$(7)\quad X_{\mathrm{r}} = \frac{Zx_{\mathrm{r}}}{f} \qquad Y_{\mathrm{r}} = \frac{Zy_{\mathrm{r}}}{f}.$$

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Finally, the 3D point in the right camera coordinate system can be transformed into the world coordinate frame by converting to homogeneous coordinates and applying a coordinate transformation ${}^{w}_{r}\mathbf{H}$

$$(8)\quad {}^{r}\boldsymbol{p} = \begin{bmatrix} X_r \\ Y_r \\ Z \end{bmatrix} \rightarrow {}^{w}\widetilde{\boldsymbol{p}} = {}^{w}_{r}\mathbf{H}\, {}^{r}\widetilde{\boldsymbol{p}}.$$

Determining the 3D world coordinates of a point ${}^{w}\boldsymbol{p}$ therefore requires us to know the camera separation, b, the focal length of the camera, f, the disparity, d, of points, and the coordinate transformation, ${}^{w}_{r}\mathbf{H}$. It is common for the camera separation and focal length of a stereo system to be known. However, it is more difficult to know the disparity and coordinate transformation than b and f.

The disparity may be found using a variety of methods. In general, there are two main categories of disparity algorithms that use sparse or dense point correspondence. Sparse methods track a limited number of features in an image and have been in development since the 1970s (Szeliski 2011), whereas dense methods compute disparities for every pixel within an image (Scharstein and Szeliski 2002). The theory we adopt here is applicable regardless of the disparity method chosen. We choose to determine the disparity using a dynamic programming solution for finding shifts in images known as smooth dynamic image warping (SDW) (Arias 2016). SDW extends dynamic image warping (DIW) (Hale 2013) by recovering sub-sample shifts. We choose the SDW algorithm because it provides dense and sub-pixel disparity values for each pixel in an image. As a dynamic programming approach to finding shifts between two images, SDW works by minimizing a constrained nonlinear optimization problem. In summary, SDW determines shift values between two images that may be used to "warp" one image to the other. In our context, we determine shifts between the left and right stereo images. Given that the cameras are only separated in the x-direction, the shift between images will only be in the x-direction. Therefore, we estimate 1D shifts between corresponding rows in left and right images. We note that estimating 2D shifts provides more constraints to the optimization problem and may be beneficial; however, we use 1D shifts in this paper for simplicity and computational efficiency. We can denote the image rows as vectors $\boldsymbol{l}$ and $\boldsymbol{r}$ where each element of $\boldsymbol{l}$, denoted l_i, is approximately an element in $\boldsymbol{r}$, r_{i+d_i}

$$(9)\quad l_i \approx r_{i+d_i} \quad \text{for} \quad i = 1, \ldots, W,$$

where W is the width of the image in pixels and d_i denotes the disparity used in eq. (6) to determine the 3D point that projects onto pixel i. We estimate the disparity for all pixels in a row, and therefore recover a dense set of 3D point coordinates using SDW. In practice, we minimize the error between $\boldsymbol{l}$ and a shifted version $\boldsymbol{l}$; SDW minimizes absolute error, which results in shifts that express how each element of $\boldsymbol{r}$ may be shifted to for a match $\boldsymbol{l}$. The details of how the SDW minimization problem is implemented are beyond the scope of this paper but can be found in (Hale 2013) and (Arias 2016). After finding the disparity using SDW, we obtain the 3D location of a point represented in the right camera coordinate frame from eqs. (6) and (7).

The cameras may be translating and rotating relative to the world coordinate frame while the stereo video is being recorded, for example, if the cameras are mounted on a hovering UAV. It is necessary to examine the 3D points in a consistent coordinate frame; therefore, we require the coordinate transformation ${}^{w}_{r}\mathbf{H}$ for each frame of the video. This coordinate transformation may be provided from an onboard IMU, which uses

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

accelerometers and gyroscopes to monitor position and orientation. In addition, other sensors, such as GPS may be used to monitor the camera motion during flight.

Many points on the ground are simultaneously viewed by the stereo cameras, and therefore we may statistically represent the position of the ground viewed by the system as a PDF. This representation of the ground assumes that between frames, all points undergo the same translation. A PDF of ground position is therefore available for each frame in the stereo video, which provides quantitative uncertainty information of the ground position (see Appendix A for detailed uncertainty calculations). The PDFs change between frames as the ground and cameras move; however, we may isolate the ground and camera motion by measuring and removing the camera motion. To remove the camera motion, we perform a coordinate transformation from right camera to world coordinate frame using eq. (1) to form ${}^{\mathrm{w}}_{\mathrm{r}}\mathbf{H}$ for a particular frame. We note that the transformation matrix ${}^{\mathrm{w}}_{\mathrm{r}}\mathbf{H}$ contains camera motion information that is independent of the ground motion (e.g., from an IMU). After camera motion is removed, we obtain a PDF from the collection of points representing the position of the ground in a consistent coordinate system as it varies with time. Summary statistics, such as the mean, may be computed from the collection of points to provide us with a "trace" that represents ground motion. Similarly, we may compute the variance of point positions as a proxy for the uncertainty of our measurements. Therefore measurements of the ground position, with associated uncertainty, are available from images of the ground taken remotely.

Numerical examples

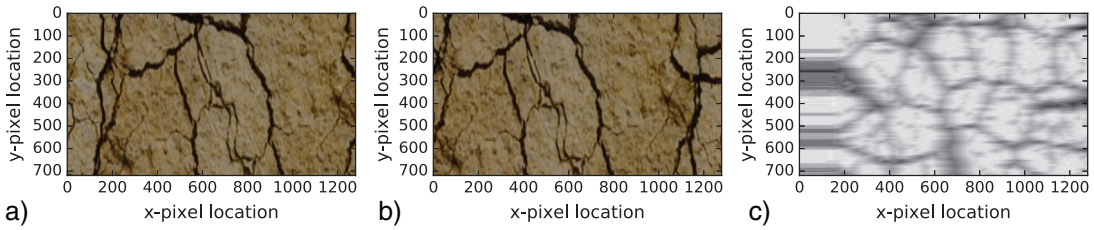
We assess the feasibility of measuring ground motion with a UAV-mounted stereo vision system using realistic virtual simulations. Virtual simulations provide flexibility and precise control of UAV and ground motion, allowing us to simulate various UAVs, cameras, and ground motion signals. By using simulations, we know true ground and camera motion, and it can be compared to the recovered motion using stereo vision. We simulate two cameras separated by $b = 30$ cm in the x-direction of the left camera frame. Each camera has a focal length of 4.15 mm, a resolution of 1280 pixels $\times$ 720 pixels, and an image sensor element size of 3.75 μm. These camera parameters are feasible and common, and were chosen to reflect current smartphone cameras. The cameras view a simulated moving ground surface from a height of 2 m, a reasonable height for a UAV hovering above the ground.

The vibrating ground surface is described by texture and topography (Fig. 2) representing cracked clay. The appearance of cracks in the images intentionally does not reflect the true cracks in the ground surface in order to show that the recovered disparities are influenced by the topography, rather than by appearance.

An earthquake signal (Fig. 3) recorded at USGS station OK034 in the fall of 2016 near Cushing, Oklahoma is used to displace the surface in three dimensions (Center for Engineering Strong Motion Data 2016). We note that the displacement of the signal is on the order of 3 mm, with many displacements below 1 mm. The vibrating ground surface is observed using the virtual stereo cameras, which hover above the ground (see Fig. 3) as if mounted on a UAV. In the following, we present recovered ground displacement results with and without drone motion.

We first show displacement recovery when the UAV is stationary. Videos are rendered for both the left and right cameras as the ground moves in three dimensions according to the ground position curves shown in Fig. 3. Dense disparity values are computed at each time step, which are used to construct a PDF representing the ground location as viewed from the stereo cameras. When estimating disparities using the SDW algorithm, we use a sub-sample precision of 1/50 and a strain range of ±0.02. The strain range chosen is low and narrow because of prior knowledge we posses: the disparities we seek are smoothly

Fig. 2. (*a*) Left, (*b*) right, and (*c*) disparity image examples for the earthquake simulation. Note the left and right images are shifted versions of one another. The shift is quantified in the disparity image on the far right; smaller (darker) disparity values are farther away from the camera and indicate cracks in the ground.



varying along a row. A wider strain range allows for more rapidly varying shifts along a row, and may be more appropriate when viewing a rugged surface. Here, the simulated ground is relatively flat, thus implying that the observed shifts vary smoothly.

The sub-sample precision chosen is sufficiently small to recover minute shifts between left and right images. We note that a higher sub-sample shift precision may be used, at the expense of an increase in computational memory requirements; however, one may not observe benefits beyond a certain sub-sample shift precision value (Arias 2016). In contrast, choosing too low of a sub-sample shift precision value may lead to poor shift recovery. We tested larger sub-sample precision values and found little benefit beyond 1/50. The sub-sample shift value is best estimated by performing disparity computations on the first frame of the video and examining the recovered shifts. If the recovered shifts vary wildly, or seem discontinuous, then a higher degree of sub-sample precision may be required. We note that if computational time and memory is not an issue, then one may set the value of sub-sample shift very high without risking the accuracy of recovered shifts. An example of recovered disparities for the first frame is shown in Fig. 2*c*. Note that the recovered disparities are somewhat smooth along a row, and they resemble cracks which reflect the cracks in the simulated ground surface.

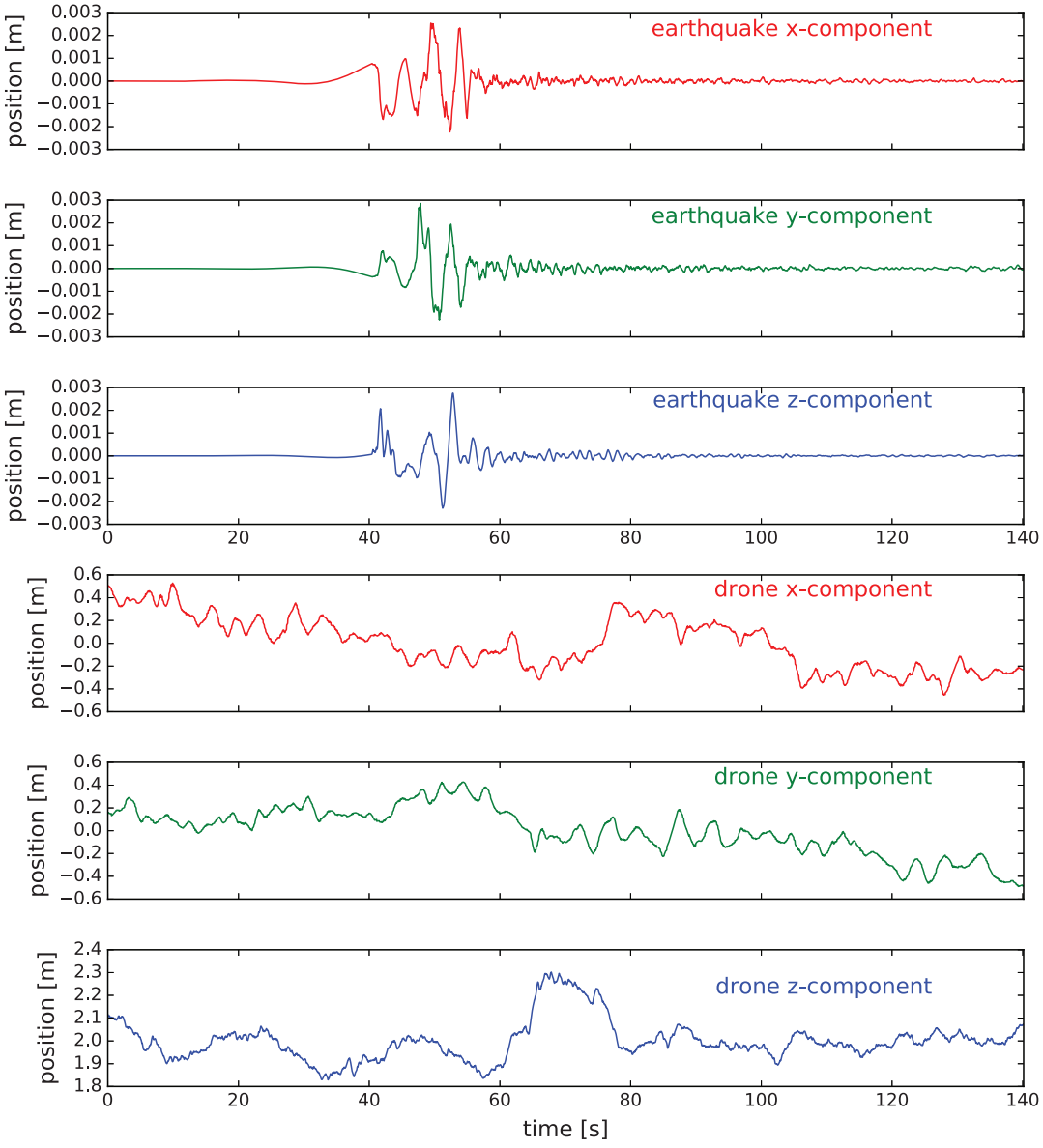
Figure 4*a* shows the recovered point location PDF in the world coordinate frame for each frame in the video. The PDF gives insight into the statistical characteristics of our measurement. The PDF is smooth, broad, and contains horizontal bands. The horizontal banding in the recovered PDF is present because of discrete disparity values recovered using SDW, which are of limited precision. Regardless, we observe in Fig. 5*a* that the recovered ground displacement, after subtracting the initial distance to the ground, traces the true displacement of the ground surface. The true and observed displacement curves are on top of one another, which shows that the signal has been recovered well. The error, for this example, has a standard deviation of 0.011 mm. The standard deviation of the error is within a hundredth of a millimetre; five orders of magnitude less than the height at which the measurement was taken ($h = 2$ m).

We repeat the experiment with a moving UAV and the same ground motion signal. When the UAV is moving, the range of disparity values we must search through using SDW is considerably larger, making disparity estimates more computationally costly. We simulate UAV motion that resembles real-life motion we measured from a DJI Matrice 100 UAV. The UAV was instructed to hover at a fixed position for about 2 min; roughly the time spanned by the earthquake signal in Fig. 3. The UAV moves a considerable amount during the simulation, roughly a quarter of a metre; this variation is orders of magnitude larger than the earthquake signal we wish to recover. The recovered PDF of point location in world coordinates for this second experiment is shown in Fig. 4*b*. We note that the banding observed in the PDF for the first experiment is not present in the PDF for the

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Fig. 3. Earthquake (top three plots) and UAV motion (bottom three plots) used in hovering UAV simulation. The red, blue, and green curves show *x*, *y*, and *z* positions. Earthquake data is recorded ground motions for the 2016 (M5.0) earthquake near Cushing, Oklahoma, at USGS station OK034. We note that the UAV motion is three orders of magnitude larger than the earthquake signal we wish to recover.



second experiment because of the coordinate transformation we use to correct for UAV motion. However, we see their remnants as scattered peaks in the recovered PDF, which seem to mimic the vertical motion of the UAV. The errors present when the UAV is moving are larger than when the UAV is stationary; we observe an error standard deviation of 0.30 mm. Based on our results, we conclude that a stereo camera can feasibly measure earthquake-like sub-millimetre ground motion. The simulation parameters we use reflect realistic UAV and camera parameters.

Fig. 4. Recovered ground position PDF (*a*) without and (*b*) with UAV motion. White curve shows the mean value of ground position, which resembles the vertical ground motion from the simulated earthquake.

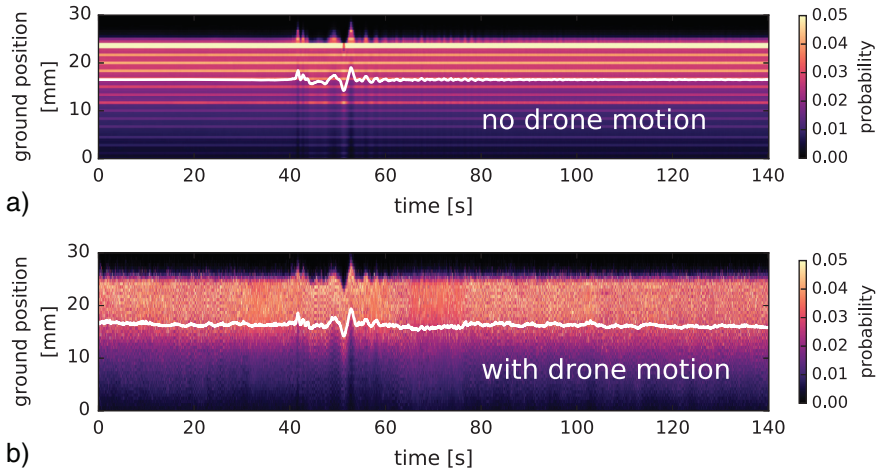
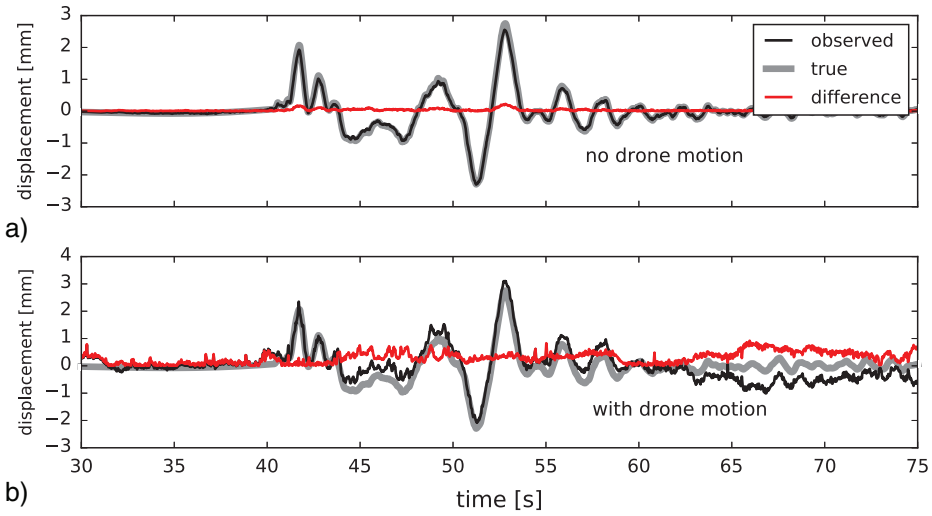


Fig. 5. Recovered ground displacement signal compared with true displacement (*a*) without and (*b*) with UAV motion. Red curves show the absolute difference between recovered and true ground motion. Drone motion influences the recovery of the input signal.



Conclusion

The UAV motion, earthquake motion, and camera specifications used in this work reflect real-life data. We assume known UAV position during acquisition, which may be available from IMUs and GPS. A dense disparity algorithm leads to many estimates of ground position at a given time, in turn allowing access to a statistical representation of ground position. We note that sub-pixel disparity estimates, provided by the disparity algorithm, are necessary to characterize the subtle earthquake ground motions. The strategy for disparity estimates used here does not provide point correspondence in time, and therefore does not allow us to recover lateral ground displacement.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

In general, current hardware and software limitations prohibit the practicality of a commercial large-scale UAV-borne exploration seismic acquisition system. The ground motion occurring during an exploration seismic survey is expected to be two to three orders of magnitude lower than the motion of an earthquake signal. Therefore, applying our current method to exploration seismic surveys will require more sophisticated camera systems with higher resolutions and larger camera offsets. In addition, the seismic exploration scenario demands a more accurate known UAV position.

In terms of hardware, it is reasonable to assume UAV and camera systems will continue to improve, allowing for more detailed images and more precise UAV positioning systems, which result in more accurate and precise ground motion measurements. In terms of software, the disparity estimate algorithm may be modified and improved. More consistent and accurate disparities may be found by including information between video frames in addition to information within a video frame. Future and ongoing work focuses on augmenting our current method to recover lateral ground displacement and rotation, and on evaluating the use of fiducial markers and a single camera system to deduce ground motion. Based on current results, we conclude that UAV-borne stereo cameras can potentially be used to measure seismic signals, with associated uncertainties.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Acknowledgements

We thank the sponsor companies of the Center for Wave Phenomena, whose support made this research possible. We thank Matt Bailey and Ross Bunker for help acquiring UAV position data. Lastly, we acknowledge using the open source computer graphics engine Blender for our videos, and appreciate those who contributed to its creation.

References

Arias, E. 2016. Seismic slope estimation: analyses in 2D/3D, M.Sc. thesis, Colorado School of Mines. Arthur Lakes Library.

Berni, A.J. 1991. Remote seismic sensing, US Patent 5,070,483.

Berni, A.J. 1994. Remote sensing of seismic vibrations by laser Doppler interferometry. Geophys. **59**(12): 1856–1867. doi: 10.1190/1.1443573.

Bostick III, F. 2000. Field experimental results of three-component fiber-optic seismic sensors. *In* SEG Technical Program Expanded Abstracts 2000. Society of Exploration Geophysicists, Tulsa, Okla. pp. 21–24.

Center for Engineering Strong Motion Data. 2016. USGS Station OK034. Available from www.strongmotioncenter.org.

Daley, T.M., Freifeld, B.M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D.E., Goetz, J., Henninges, J., and Lueth, S. 2013. Field testing of fiber-optic distributed acoustic sensing (das) for subsurface seismic monitoring. The Leading Edge, **32**(6): 699–706. doi: 10.1190/tle32060699.1.

Hale, D. 2013. Dynamic warping of seismic images. Geophys. **78**(2): S105–S115. doi: 10.1190/geo2012-0327.1.

Mateeva, A., Mestayer, J., Cox, B., Kiyashchenko, D., Wills, P., Lopez, J., Grandi, S., Hornman, K., Lumens, P., Franzen, A., Hill, D., and Roy, J. 2012. Advances in distributed acoustic sensing (DAS) for VSP. *In* SEG Technical Program Expanded Abstracts 2012. Society of Exploration Geophysicists, Tulsa, Okla. pp. 1–5.

Mateeva, A., Lopez, J., Mestayer, J., Wills, P., Cox, B., Kiyashchenko, D., Yang, Z., Berlang, W., Detomo, R., and Grandi, S. 2013. Distributed acoustic sensing for reservoir monitoring with VSP. The Leading Edge, **32**(10): 1278–1283. doi: 10.1190/tle32101278.1.

Mateeva, A., Lopez, J., Potters, H., Mestayer, J., Cox, B., Kiyashchenko, D., Wills, P., Grandi, S., Hornman, K., Kuvshinov, B., Berlang, W., Yang, Z., and Detomo, R. 2014. Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling. Geophys. Prospect. **62**(4): 679–692. doi: 10.1111/1365-2478.12116.

Mestayer, J., Cox, B., Wills, P., Kiyashchenko, D., Lopez, J., Costello, M., Bourne, S., Ugueto, G., Lupton, R., Solano, G., Hill, D., and Lewis, A. 2011. Field trials of distributed acoustic sensing for geophysical monitoring. *In* SEG Technical Program Expanded Abstracts 2011. Society of Exploration Geophysicists, Tulsa, Okla. pp. 4253–4257.

Quam, L.H., Tucker, R.B., Liebes, Jr, S., Hannah, M.J., and Eross, B.G. 1972. Computer interactive picture processing, Technical Report 166, Stanford Artificial Intelligence Project Memo.

Scharstein, D., and Szeliski, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Comput. Vis. **47**(1–3): 7–42. doi: 10.1023/A:1014573219977.

Stanbridge, A., Ewins, D., and Khan, A. 2000. Modal testing using impact excitation and a scanning ldv. Shock Vib. **7**(2): 91–100. doi: 10.1155/2000/527389.

Szeliski, R. 2011. Computer vision: algorithms and applications. Springer.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.

Appendix A: Analytic uncertainty in depth estimates

It is common and beneficial to analytically determine how error in projected pixel locations translates into disparity error. We may quantify the change in disparity given a small change, or error, in the pixel locations Δx_r or Δx_l using total derivatives, which for a function $z = f(x, y)$ can be approximated as

$$\Delta z \approx \frac{\delta f}{\delta x}\Delta x + \frac{\delta f}{\delta y}\Delta y, \tag{A1}$$

where small changes in x and y are denoted as Δx and Δy, respectively. As discussed in the manuscript, the disparity is related to the focal length (f), the camera separation (b), and the depth from camera (z). We can then compute the error in disparity $d(x_r, x_l) = fb/z$ as

$$\Delta d = \frac{\delta d}{\delta x_r}\Delta x_r + \frac{\delta d}{\delta x_l}\Delta x_l = \Delta x_r - \Delta x_l, \tag{A2}$$

where Δx_r and Δx_l are errors in the point projection on the right and left cameras, respectively. We expect the errors in the left and right cameras to be independent, as they were taken from independent cameras, and assume the errors are zero mean. The expected value of the disparity error then becomes

$$\mu_d = E[\Delta d] = E[\Delta x_r - \Delta x_l] = E[\Delta x_r] - E[\Delta x_l] = 0. \tag{A3}$$

The variance of the disparity error can be computed by leveraging eq. (A3) and using the pixel error independence assumption to obtain

$$\begin{aligned}
\sigma_d^2 &= E[(\Delta d - \mu_d)^2] \\
&= E[\Delta d^2] \\
&= E[(\Delta x_r - \Delta x_l)^2] \\
&= E[\Delta x_r^2] + E[\Delta x_l^2] - 2E[\Delta x_r \Delta x_l] \\
&= \sigma_l^2 + \sigma_r^2 - 2\sigma_{lr}^2 \\
&= \sigma_l^2 + \sigma_r^2.
\end{aligned} \tag{A4}$$

We conclude that the mean of the disparity error is zero given independent and zero mean errors in the left and right camera pixel projection coordinates. The variance of the disparity is then the sum of the variance of the errors in the left and right cameras.

We repeat this analysis to derive the depth estimate error given an error in disparity

$$\Delta Z = \frac{\delta Z}{\delta d}\Delta d = -\frac{fb}{d^2}\Delta d. \tag{A5}$$

We see that the absolute value of the distance error is proportional to fb/d^2. Proportionally, this result implies that an increase in camera separation (b) or focal length (f) increases the distance error for a given disparity (d). However, the disparity dominates the depth error as $1/d^2$, implying that large disparities can drastically decrease the depth error. The mean of the error in depth Z can be found by leveraging eq. (A5) and simplifying the ΔZ term,

$$\mu_Z = E[\Delta Z] = -\frac{fb}{d^2}E[\Delta d] = 0, \tag{A6}$$

while the variance of the error in depth Z is,

$$\begin{aligned}\sigma_Z^2 &= E[(\Delta Z - \mu_Z)^2] \\ &= E[\Delta Z^2] \\ &= E\left[\left(-\frac{fb}{d^2}\right)^2 \Delta d^2\right] \\ &= \left(\frac{fb}{d^2}\right)^2 \sigma_d^2.\end{aligned} \tag{A7}$$

The standard deviation of the depth error is proportional to the quantity $(fb/d^2)^2$. This quantity is dominated by the effects of d^{-4}; if the disparity increases, then the variance in the depth error rapidly decreases. While mean and variance summarize the behavior of the distance error, a more complete representation of the displacement is represented using PDFs, as demonstrated in the manuscript. Based on this analysis, we expect the error to deteriorate when disparities are small (i.e., when the UAV is farther away from the ground). Thus, the displacement error we obtain when the UAV is moving depends on the UAV motion.

J. Unmanned Veh. Sys. Downloaded from cdnsciencepub.com by Thomas Rapstine on 12/16/21
For personal use only.