Journal of Electronic Imaging

JElectronicImaging.org

Motion characterization using optical flow and fractal complexity

Joshua D. Borneman
Evie Malaia
Ronnie B. Wilbur

Joshua D. Borneman, Evie Malaia, Ronnie B. Wilbur, "Motion characterization using optical flow and fractal complexity," *J. Electron. Imaging* **27**(5), 051229 (2018), doi: 10.1117/1.JEI.27.5.051229.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Motion characterization using optical flow and fractal complexity

Joshua D. Borneman,[a,*] Evie Malaia,[b,c] and Ronnie B. Wilbur[a]
[a]Purdue University, West Lafayette, Indiana, United States
[b]Freiburg Institute for Advanced Studies, Freiburg im Breisgau, Germany
[c]University of Alabama, Department of Communicative Disorders, Tuscaloosa, Alabama, United States

Abstract. We developed a technique, using fractal complexity analysis of optical flow in two-dimensional (2-D) videos, to characterize information content in observed motion. Several lines of evidence demonstrate that visually available properties of motion can characterize the state of a system. This paper will describe the method used and will present a test case regarding the accuracy of the method. An analytical comparison of simple human movement (arranging items on a table) and American Sign Language (ASL) will be given as an example application. The normalized spectral density in the range of 0.1 to 15 Hz indicated significantly higher fractal complexity in the optical flow of ASL video data, indicating that information content in 2-D video data can be characterized using complexity analysis of optical flow. The technique used for quantification of information content in visual motion data is likely to be applicable for distinguishing biological versus nonbiological motion in 2-D video data, making inferences about the states of biological objects from the dynamics of optical flow, and in assessing likelihood of information content in a video stream. *© 2018 SPIE and IS&T* [DOI: 10.1117/1.JEI.27.5.051229]

Keywords: optical flow; fractal complexity; communication; information transfer; sign language; biological motion.

Paper 171108SS received Dec. 20, 2017; accepted for publication Jun. 12, 2018; published online Jul. 11, 2018.

1 Introduction

The problem of automated recognition of biological motion is a crucial problem in many applications, from automated surveillance to human–robot interaction.[1–3] Currently, human activity recognition typically relies on feature encoding, such as histogram of oriented optical flows (HOOF)[4,5] and other methods.[6–9] However, these methods are based on empirical analysis, and utilize no a priori knowledge of the system being observed.

Human motion however, as a biological function, is based on interaction of multiple neural and motor systems at a range of temporal and spatial scales, which can characterize the state of a living system (e.g., healthy participant versus Parkinson's patient[10]). In the field of complex systems biology, such multiscale systems are typically described using entropy analysis. Entropy-based measures relate to the degree of the signal's regularity or predictability over time and describe the quantitative complexity of the system generating the signal, and well as potential information throughput within the signal.[11]

The techniques for estimating entropy of the signal are derived from the field of nonlinear dynamics; these techniques aim to quantify variability and correlated properties in dynamic signals, which requires mathematical representations of the signal by time series. Linear output of multiscale systems (e.g., motion capture data and neural spikes spectra) typically exhibits variability that can be described by power laws. However, multidimensional recordings of biological systems, and video recordings of human motion, present a computational challenge with respect to their reliability and robustness. In the present work, we demonstrate a method of entropy analysis on biological data from recorded videos.

1.1 *Biological Basis*

Studies of human processing of auditory[12] and visual motion[13–15] have begun to investigate the human capabilities to parse and interpret biological signals. Early research in event segmentation has identified neural activity time-locked to velocity changes in the scenes,[16] and sign language users were shown to be sensitive to smaller differences in acceleration patterns, as compared with non-signers,[17] likely due to experience with higher information density in the visual signal.[18,19] Humans from unrelated cultures can identify emotional states of the moving person based on the profile of biological motion; both signers and nonsigners are able to successfully extract linguistic aspectual meaning from hand motion.[14,20,21] The algorithms for information extraction from visual motion in the human brain are not yet clear; however, humans still surpass state-of-the-art computer vision approaches.[22] Prior work shows an indication that human visual information transfer relies on recognition of velocity, acceleration patterns, and recognition of complex (as in communication theory) motion.

The use of entropy analysis of video data opens new venues for automatic systems aimed at filtering such data for information transfer, or for providing assistance to human operators tasked with identification motion in video recordings.

2 Method

To extract the information entropy,[11] or fractal complexity, of a given type of motion, an initial video collection is obtained, which contains the relevant type of motion.

The following motion analysis was conducted using custom scripts within MathWorks, Inc. MATLAB. The analysis presented here was originally run on videos of a person either

*Address all correspondence to: Joshua D. Borneman, E-mail: joshua.borneman@gmail.com

1017-9909/2018/$25.00 © 2018 SPIE and IS&T

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

signing in American Sign Language (ASL) or performing normal everyday activities.[19]

At a high level, the videos are preprocessed to achieve uniform object size and resolution (Sec. 2.2). Optical flow is then calculated and an optical flow histogram (a velocity spectrum) is created for each single video frame (Sec. 2.3). A temporal Fourier transform is then used on the velocity spectrum versus time to obtain a power spectral density per velocity (Sec. 2.3). This frequency dependence of the power spectral density (PSD) function shows the fractal complexity of the information transfer due to the motion in the video. The fractal complexity parameter is extracted from the power spectral density of motion (Sec. 2.4).

2.1 *Source Video Temporal Preprocessing*

The original source video datasets contain videos of different lengths. The frame rate and duration of each video determine the temporal frequencies that can be calculated, so these parameters will bound the sampling of the extracted PSD function. The frequency resolution (f) is $f_{\max} = r/2$ and $f_{\text{step}} = 1/T$, where $f_{\max}$ (Hz) is the maximum-resolved frequency, r (Hz) is the frame rate of the video, f_{step} (Hz) is the temporal frequency resolution, and T (seconds) is the full video duration. The test case of 30 frames per second video results in Frequencies defined from 0.01 to 15 Hz in 0.01 Hz steps.

Given a small number of long videos in the dataset, they may be cropped into individual data samples, provided that T is sufficiently long, and that the content at later times may be analyzed as an independent dataset. For the sign language example,[19] each video clip of 45 seconds is long enough to contain 40 to 50 sentences, and each clip contains similar information transfer, but a different specific content. To achieve common temporal frequency profiles in the final analysis, all input videos should have common frame rate, and should be cropped to a consistent duration.

2.2 *Source Video Spatial Preprocessing*

Processing will rely on an analysis of the global motion in the video, and so videos with a static background, and in which only the motion in question is present, are ideal. Video sets with camera motion may require an initial preprocessing step to stabilize the video according to background objects and may be sufficient to isolate the information-containing motion.[23,24] However, our example analysis uses videos from prior experiments, which contained a static background, so the utility of stabilization was not examined further.

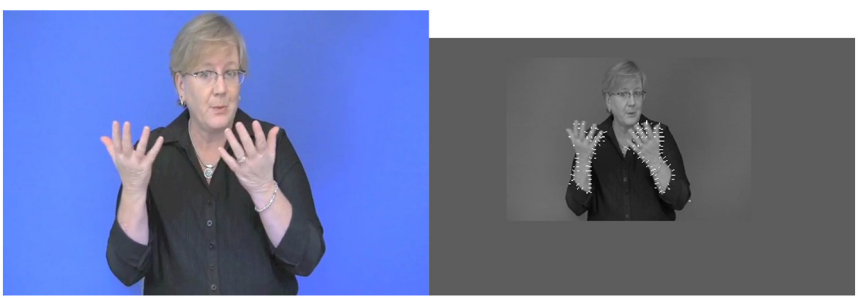
All videos were converted to an eight-bit grayscale as the analysis performed in this paper is concerned only with motion.

Optical flow of a video frame is the distribution of apparent velocities of objects in an image; that is, a velocity vector (in pixels/frame) is found for each pixel, based on how fast and in which direction, the feature shown in that pixel has moved from the frame before. The collected dataset videos may contain different fields of view and distances to the source of motion (the person, in our work), which were not quantitatively known. In this case, each pixel represents a different sampling size on the person, and therefore differences in optical flow values (pixels/frame) might be found due only to differences in the setup of each camera.

To eliminate differences in motion velocity magnitude across the dataset, the videos must be scaled to achieve uniform person size, and consistent resolution. Appropriate scaling requires a common reference object on all videos, after which the videos were resized so that this reference was the same length (in pixels) across the entire video dataset. In our analysis, reference selection was done with manual user input, where the point of the shoulder and point of the elbow were manually selected. Reference object length, in pixels was then calculated. A scale factor for each individual video was determined as $D_{\min}/D_n$, where $D_{\min}$ is the global minimum dimension of the reference across all videos, and D_n is the dimension of the reference on the n'th video. This results in a modified video set where the reference object is equal to $D_{\min}$ across the entire video set.

Analyses were also done with unscaled videos, which determined that the relative changes in velocity, or the frequency components of the optical flow, actually remained unchanged with scaling; however, scaling the videos allows us to obtain similar values for optical flow velocity magnitude across all videos for each step of the analysis.

Scaling for a consistent reference object size results in an object resolution equal to the least-common resolution of the full video set, $D_{\min}$, and results in videos with different total resolutions, even if the original video dataset had a uniform resolution. Therefore after scaling, the videos were padded and/or cropped so that all videos were the same final resolution. Uniform resolution creates a consistent spatial-frequency sampling of the object of motion (person) across the entire dataset. When cropping, the maximum object distance away from the center of the video frame, in pixels,



(a) (b)

Fig. 1 Frame from original sign language video, (a) color 720 × 480, and the same frame after preprocessing all videos, (b) grayscale 500 × 301. Optical Flow magnitude vectors are shown in white.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

was identified for each video as n horizontal and m vertical pixels. The maximum number of pixels over the dataset, $(2 \cdot n_{\max}, 2 \cdot m_{\max}) = (N, M)$, sets crop size for all videos. No motion information was lost; the object (the person's hands) never left the final frame on any video in the dataset. Where padding was needed to match the dimensions of the entire dataset, a single-color border (gray value equal to the average of the entire video frame) was added to the video frame. An example of original, and processed, frame is shown in Fig. 1.

At this point, every video in the preprocessed video dataset has dimensions $N \times M$, duration T, and contain a common reference object of length $D_{\min}$.

2.3 *Optical Flow*

Optical flow (OF) for each video was determined using the MathWorks' MATLAB vision toolbox optical flow function. This function was utilized to compare each video frame with the prior frame and, using a Horn–Schunck method,[25] an output matrix of size equal to the input video frame was calculated. Each element of the matrix identifies the magnitude of optical flow velocity (pixels per frame) between the two frames for each corresponding pixel in the video.

Although a region of interest (ROI) could be used in cluttered videos to analyze the motion of only one object, our test videos had a uniform background with only the object of interest in motion, therefore we did not preserve the spatial (x, y) origins of each optical flow vector, only the overall motion profile. Therefore, the OF matrix (sized $N \times M$) for each frame of video is reshaped so that it is characterized by a single vector of OF magnitudes, $A(i)$, where i is the position with $(N \cdot M)$ elements. Collected over the entire video, we obtain a matrix $\mathbf{A}(i, t)$, where t is the frame time from 0 to T. A is then binned into a histogram $B(j, t)$, where j is one of the J optical flow velocity bins. Values of 0, no motion, are not collected. The highest velocity bin value is determined by initially processing the entire video set to determine the maximum OF (pixels/frame) from the entire video dataset. OF values could be converted to a real-world velocity value if camera parameters are known. With the angular subtend of each pixel (radians) and the frame rate, optical flow could be converted to radians/sec. However for the analysis here, it is sufficient to have all videos scaled identically, and to remain in video coordinates (pixels/frame). An example optical flow histogram versus time, B, is shown in Fig. 2.

For each optical flow velocity bin, we look at the changes to that velocity component over time, i.e., at the frequency modulation of the optical flow signal. The power spectral density (PSD) was calculated using a MATLAB's "pwelch" PSD estimate. The PSD is taken for each bin j, resulting in J separate frequency profiles so that $M[j, f]$ has size J (bins) by the resolved frequency vector f, as described in Sec. 2.1. This gives a vector for the power spectral density of the motion (optical flow) versus frequency. This is shown for a sign language video in Fig. 2.

2.4 *Fractal Complexity*

The optical flow power spectral density was then analyzed according to its fractal complexity.[11] The function given in Eq. (1) was fit to each frequency profile, where M is the magnitude of optical flow, f is the frequency, α is the spectral density amplitude, and β is the fractal complexity. The exact method used is described in more detail in Sec. 2.5, as method 2, the linear log/log fit method

$$M(j, f) = \frac{\alpha}{|f|^{\beta}}. \tag{1}$$

Thus, for each of the j optical flow velocity bins, an amplitude fitting variable and the fractal complexity are found, $\beta(j)$. Additionally, $M(j, f)$ is integrated over j, obtaining an overall velocity spectrum for the video, and after fitting, an overall fractal complexity $\bar{\beta}$ is found, which represents the amount of information transfer, or fractal complexity, for the motion in the video.

2.5 *Fractal Complexity Accuracy*

Significant work was done to validate the method and accuracy of extracting $\bar{\beta}$ through a best-fit from the $M(f)$ optical flow. As a test, a simulated noise signal was created with a defined fractal complexity. Various extraction methods were tested to retrieve the a priori value, and to quantify the error of the method.

Amplitude noise signals, $V(t)$, with 2^{14} elements and various defined fractal complexity (i.e., noise color) β,

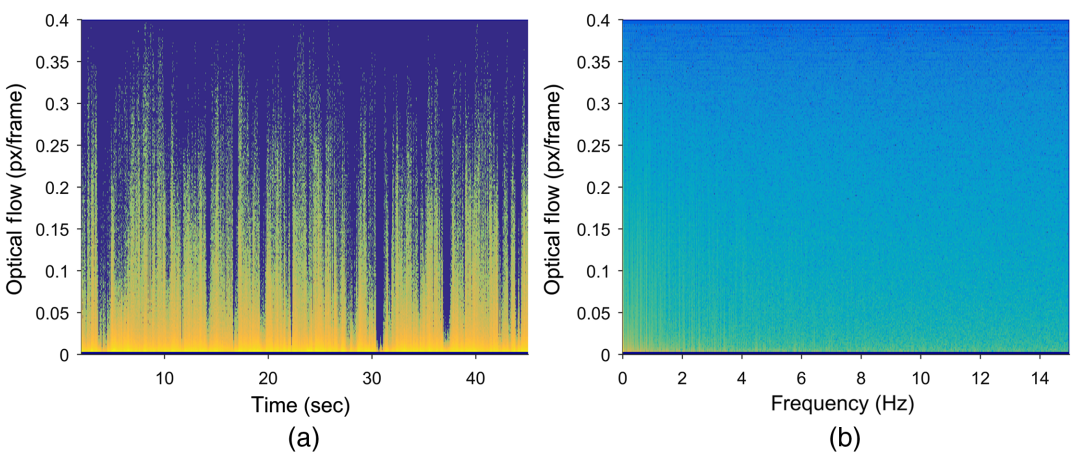


Fig. 2 (a) Velocity spectrum versus time and (b) the power spectral density per OF for a sign language video (45 s, 30 fps).

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

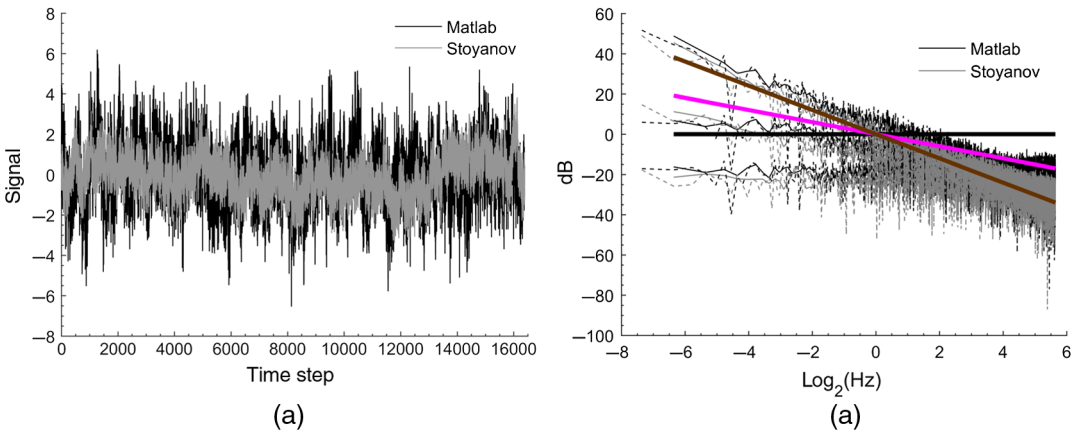


Fig. 3 (a) Test signals ($\beta = 1$) generated using the default MATLAB method,[26] and Stoyanov[27] methods. (B) PSD of test signals ($\beta = 0$, 1, and 2), plotted on a dB per octave scale. Solid lines are PSD using pwelch and dashed lines are PSD using a manual fft. Thick lines are reference lines for $\beta = 0$ (white), 1 (pink), and 2 (brown) noise.

were generated so that the PSD of $V(t)$ should be accurately described by Eq. (1). Noise was generated using both the default MATLAB method (dsp. ColorNoise) function[26] and another published method,[27] i.e., "Stoyanov" method.

Then the PSD of $V(t)$, $M(f)$, was found using both the "pwelch" function, as in Sec. 2.3, and using a manual FFT script. Example test signals and the PSD of those signals are shown in Fig. 3. Both methods of PSD calculation agree quite well, Fig. 3(b). Deltas may be due to differences in frequency sampling based on the methods. Particularly, Hamming window size for pwelch changes the frequency sampling of the result.

Last, two methods were used to re-extract the fractal complexity, or color, from the test signal. Method 1, an inverse fit method; where the inverse function, Eq. (1), is fit directly using a Matlab nonlinear least squares fit to $M(f)$ using α and β as fitting parameters.

Method 2, linear log/log method; where Eq. (1) is approximated as shown in Eq. (2). A simple linear fit is then performed on $\ln(f)$ versus $\ln(M)$, where β is the slope and α is the intercept on a log–log plot

$$\ln(M) \cong \beta \cdot \ln(f) + \alpha. \tag{2}$$

The final accuracy of the extracted β is examined by comparing it with the original signal generation parameter, using the ratio $\beta_{\text{extracted}}/\beta_{\text{generated}}$, where a ratio 1 is a perfect fractal complexity extraction.

For method 2, the linear log/log method, the extracted β of the MATLAB dsp noise is 0.91 of the intended value. For large Hamming Windows, the extracted β of the Stoyanov noise is 0.90 of the intended value. Method 1, direct inverse function fitting, results in an unacceptable and variable error across the range of expected β values.

Residual error in the log/log fit method could be due to either error in the noise generation, meaning that the extracted β is correct, or error in PSD and β fitting. Figure 4 shows that the error in β is a constant value (0.90) across any reasonable range [0 to 2] of fractal complexity for either a one-dimensional 1-D (time series) or for a 2-D (image),[28] and contains the estimated value of entropy for the English language, 1.3.[29] Therefore, while the absolute fractal complexity extracted from a signal may have roughly

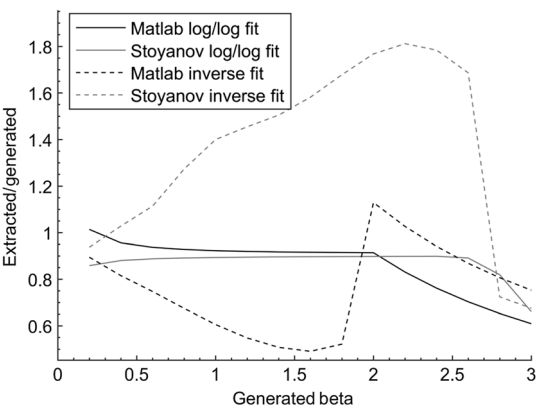


Fig. 4 Ratio of $\beta_{\text{extracted}}/\beta_{\text{generated}}$ for various $\beta_{\text{generated}}$, for MATLAB and Stoyanov noise, extracted using the linear log/log fit method [Eq. (2)] and the inverse function fit method [Eq. (1)].

10% error, the comparison of relative fractal complexity (the ratio comparing one video set to another) will be correct. Therefore, comparing the fractal complexity, or information transfer, between two video datasets should not have significant error.

3 Results for Human Motion and Sign Language Videos

Videos for both everyday motion[30] and sign language[13,31] were provided for this analysis from prior studies on motion-event boundaries and ASL predicates, respectively. Detailed results comparing everyday motion to sign language were discussed previously.[19]

Videos contained a participant in front of a static, uniform background. The signing and nonsigning videos (20 of signing and 40 of everyday motion) contained 1350 frames (30 fps, 45 s) and had been recorded at 768×512 pixels.

In the test video set, the upper arm length was selected because it was most often perpendicular to the camera axis. Potential references such as hands and forearms were considered but proved problematic due to rotation directly toward or away from the camera, resulting in an

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

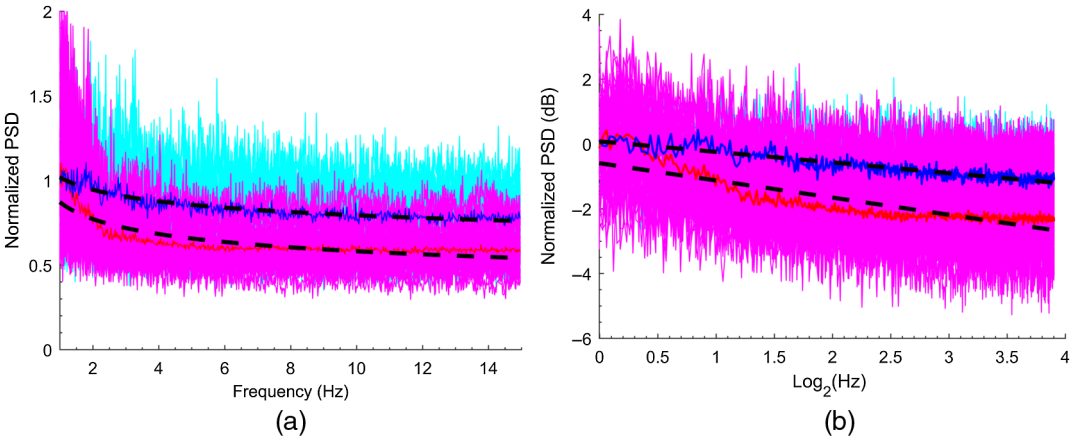


Fig. 5 Normalized spectral density for ASL (cyan/blue) and everyday motion (magenta/red). Black lines show the respective fit. Signing videos show greater fractal complexity.[19] (a) Linear scale and (b) dB per octave.

artificial reduction in size on the video frame, which would affect proper scaling.

The resulting final dataset of 60 videos was all 500×301, grayscale, 30 fps, and 45 s duration, with a consistent participant size. Frequencies are defined from 0.01 to 15 Hz based on the video duration (45 s) and frame rates (30 Hz). Optical flow was binned into 200 bins from 0 to 0.4 pixels per frame.

Analysis of these videos has shown that spectral density amplitude and fractal complexity parameter were extracted with an average root mean square error (RMSE), across all fits, of 2%, with a global maximum RMSE of 3%. The extracted fractal complexity for a subset is shown in Fig. 5, also showing normalized fractal complexity versus frequency.

Using the method proposed in this paper, signing videos show greater fractal complexity than simple human motion.

4 Discussion

This paper proposes a method to classify motion in a video using a priori knowledge of motion of a complex system (such as a biological system) and based on theories of how humans detect and classify observed motion.[32–34] We demonstrate a method of entropy analysis on biological data from recorded videos.

Any human motion is, to some degree, informative to humans. Multiple studies have indicated that participants automatically and with high degree of precision segment both everyday motion and folk dancing,[30,35] as well as extract basic event-level information from visual motion in unfamiliar sign languages.[14,36] Information throughput in sign language videos was assumed to be higher for several reasons. First, in sign language data, the communicative goal—information transfer—was implicit: the videos used were informative and comprehensible to native ASL signers. Thus, while both simple biological and sign language-related motion both conform to power law, the latter was shown to have higher entropy, or capacity for carrying information.[37] Our analysis has shown that the sign language motion had a greater fractal complexity, and thus, the video had a higher potential information-carrying capacity.

Generally, a technique to quantify the information content in visual motion data has been presented. This technique may be extended to other types of motion classification than those discussed here. For example, the fractal complexity of some motion (human walking, human with a limp, or human with Parkinson's) may be detectible, although higher quality recordings may be necessary. We propose that the fractal complexity of the optical flow time series is a useful parameter to aid in the detection and classification of human motion. It is robust to different video input resolutions and scaling, and it should be independent of aspect. The functional techniques used have been tested against simulated datasets and the information extraction (fractal complexity) absolute error of this method has been quantified and found to be <10%, whereas the relative fractal complexity is well behaved across datasets, and value comparisons should be accurate. This approach is likely to be applicable for distinguishing biological versus nonbiological motion in 2-D video data, making inferences about the states of biological objects from the dynamics of optical flow, and in assessing likelihood of information content in a video stream.

References

1. J. Sattar and G. Dudek, "Visual identification of biological motion for underwater human-robot interaction," *Auton. Robots* **42**(1), 111–124 (2018).
2. A. Ladjailia et al., "Encoding human motion for automated activity recognition in surveillance applications," in *Computer Vision: Concepts, Methodologies, Tools, and Applications*, pp. 2042–2064, Information Resources Management Association (2018).
3. T. Ko, "A survey of behavior analysis in video surveillance for homeland security applications," in *Applied Imagery Pattern Recognition Workshop*, Washington DC (2008).
4. R. Chaudhry et al., "Histograms of oriented optical flow and Binet–Cauchy kernels on nonlinear dynamical systems for the recognition of human actions," in *IEEE Conf. on Computer Vision and Pattern Recognition* (2009).
5. B. C. Ustundag and M. Unel, "Human action recognition using histograms of oriented optical flows from depth," in *Advances in Visual Computing. ISVC 2014*, G. Bebis et al., Eds., Vol. **8887**, pp. 629–638, Springer, Cham (2014).
6. A. Fod, M. J. Mataric, and O. C. Jenkins, "Automated derivation of primitives for movement classification," **12**(1), 39–54 (2002).
7. A. G. Hochuli et al., "Detection and classification of human movements in video scenes," in *Advances in Image and Video Technology. PSIVT 2007*, D. Mery and L. Rueda, Eds., Vol. **4872**, Springer, Berlin, Heidelberg (2007).

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

8. S. Zhang et al., "A review on human activity recognition using vision-based method," *J. Healthcare Eng.* **2017**, 3090343 (2017).
9. F. S. Khan and S. A. Baset, "Real-time human motion detection and classification," in *Proc. IEEE Students Conf.* (2002).
10. M. Sekine et al., "Fractal dynamics of body motion in patients wtih Parkinson's disease," *J. Neural Eng.* **1**(1), 8–15 (2004).
11. C. E. Shannon, "Mathematical theory of communication," *Bell Syst. Tech. J.* **27**, 379–423 (1948).
12. N. C. Singh and F. E. Theunissen, "Modulation spectra of natural sounds and ethological theories of auditory processing," *J. Acoust. Soc. Am.* **114**(6), 3394–3411 (2003).
13. E. Malaia et al., "Event segmentation in a visual language: neural bases of processing American Sign Language predicates," *Neuroimage* **59**(4), 4094–4101 (2012).
14. B. Strickland et al., "Event representations constrain the structure of language: sign language as a window into universally accessible linguistic biases," *Proc. Natl. Acad. Sci. U. S. A.* **112**(19), 5968–5973 (2015).
15. L. A. Petitto, S. Holowka, and L. E. Sergio, "Language rythms in baby hand movements," *Nature* **413**, 35–36 (2001).
16. J. M. Zacks et al., "Human brain activity time-locked to perceptual event boundaries," *Nat. Neurosci.* **4**(6), 651–655 (2001).
17. E. S. Klima et al., "From sign to script: effects of linguistic experience on perceptual categorization," *J. Chin. Linguist. Monagr. Ser.* (13), 96–129 (1999).
18. E. Malaia, J. D. Borneman, and R. B. Wilbur, "Information transfer capacity of articulators in American Sign Language," *Lang. Speech* **61**, 97–112 (2018).
19. E. Malaia, J. D. Borneman, and R. B. Wilbur, "Assessment of information content in visual signal: analysis of optical flow fractal complexity," *Visual Cognit.* **24**, 246–251 (2016).
20. E. Malaia and R. B. Wilbur, "Kinematic signatures of telic and atelic events in ASL predicates," *Lang. Speech* **55**(3), 407–421 (2012).
21. E. Malaia, R. B. Wilbur, and M. Milkovic, "Kinematic parameters of signed verbs," *J. Speech Lang. Hear. Res.* **56**(5), 1677–1688 (2013).
22. A. Barbu et al., "Seeing is worse than believing: reading people's minds better than computer-vision methods recognize actions," in *European Conf. on Computer Vision*, pp. 612–627 (2014).
23. A. Litvin, J. Konrad, and W. C. Karl, "Probabilistic video stabilization using Kalman filtering and mosaicking," in *IS&T/SPIE Symp. on Electronic Imaging, Image and Video Communications and Proc.* (2003).
24. Y. Matsushita et al., "Full-frame video stabilization," in *Computer Vision and Pattern Recognition* (2005).
25. B. K. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.* **17**, 185–203 (1981).
26. J. Kasdin, "Discrete simulation of colored noise and stochastic processes and 1/f^a power law noise generation," *Proc. IEEE* **83**(5), 802–827 (1995).
27. M. Stoyanov, "Color noise," 2015, ~http://people.sc.fsu.edu/~jburkardt/m_src/cnoise/cnoise.html
28. B. B. Mandelbrot, *Fractals: Form, Chance and Dimension*, Librarie du Bassin (1977).
29. B. Schneier, *Applied Cryptography*, 2nd ed., John Wiley and Sons (2017).
30. J. M. Zacks et al., "Using movement and intentions to understand human activity," *Cognition* **112**(2), 201–216 (2009).
31. E. Malaia and R. B. Wilbur, "Telicity expression in the visual modality," in *Telicity, Change, and State: A Cross-Categorial View of Event Structure*, pp. 122–136 (2012).
32. R. B. Wilbur and E. Malaia, "Contributions of sign language research to gesture understanding: what can multimodal computational systems learn from sign language research," *Int. J. Semantic Comput.* **2**(1), 5–19 (2008).
33. E. Malaia, "Current and future methodologies for quantitative analysis of information transfer in sign language and gesture data," *Behav. Brain Sci.* **40**, e63 (2017).
34. E. Malaia and R. B. Wilbur, "Sign languages: contribution to neurolinguistics from cross-modal research," *Lingua* **120**(12), 2704–2706 (2010).
35. K. Noble et al., "Event segmentation and biological motion perception in watching dance," *Art Percept.* **2**(1), 59–74 (2014).
36. J. Fenlon et al., "Seeing sentence boundaries," *Sign Lang. Linguist.* **10**(2), 177–200 (2008).
37. E. Malaia, "It still isn't over: event boundaries in language and perception," *Lang. Linguist. Compass* **8**(3), 89–98 (2014).

Joshua D. Borneman is the chief scientist for Electro-Optic Technology at Naval Surface Warfare Center Crane and is also a visiting scholar at Purdue University. He received his MS degree in physics from Purdue University in 2004, and his PhD in optical metamaterials from Purdue University in 2010. He is the author of 32 journal papers and conference presentations. His interests cover advanced electro-optic and laser systems, optical metamaterials, and image signal processing.

Evie Malaia is an associate professor in the Department of Communicative Disorders at the University of Alabama, Tuscaloosa, and a Marie Curie senior fellow at the Freiburg Institute for Advanced Studies, Germany, from 2017 to 2018. She received her PhD in linguistics from Purdue University in 2005. Her research program focuses on computational modeling of communication signal (visual and auditory), and analysis of its neural processing in typical and atypical populations.

Ronnie B. Wilbur is a professor of linguistics and a professor of speech, language, and hearing sciences at Purdue University. She received her PhD in linguistics from the University of Illinois. She has conducted groundbreaking research on sign languages, including syllable and word formation, grammar and semantics, and their neurolinguistics processing. A central focus of this work is the application to effective education of deaf children, as sign language proficiency is the single predictor variable of academic achievement.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 7/11/2018
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use