Research Article

MATHEMATICALLY OPTIMAL RESTORATION AND STABILIZATION OF THE PATTERSON-GIMLIN FILM WITH COMPUTATION FEATURE DETECTION

Isaac Y. Tian[1], Bill Munns[2], Jeff Meldrum[3]

[1] *Paul G. Allen School of Computer Science & Engineering, University of Washington, 3800 E Stevens Way NE, Seattle, WA 98195-2355*

[2] *Blue Jay, CA, 92317*

[3] *Department of Biological Sciences, Idaho State University, 921 S. 8th Ave, Pocatello, ID 83209*

ABSTRACT. The Patterson-Gimlin Film (PGF) to date remains the clearest purported video evidence of the unrecognized bipedal primate known as "sasquatch". Previous analysis of the (PGF) relied on manual manipulation of individual frames to stabilize the erratic camera motion and eliminate image noise and copying defects. We utilize modern computer vision algorithms and a large multi-copy film scan database to generate a mathematically optimal frame stabilization sequence with the clearest image quality restored to date. Eight independent prints of 150 frames of the PGF surrounding the lookback frame at Frame 354 (F354) photographed at close range with a 12-megapixel (MP) digital single lens reflex (DSLR) camera and were computationally aligned and merged per frame index with the SIFT algorithm (Lowe, 2004). The composited frames were then aligned to a stationary background scene from F354 using a 3D homography solved for with the RANSAC algorithm (Fischler, 1981). All composited frames were color corrected to the background scene by solving a linear regression per color channel between the composite frame and the background frame. The resulting composited frame sequence contained more original image detail and less artifacts than any individual copy. Our rendition is an unbiased machine optimized solution that is not susceptible to injected features from manual photo editing or neural net interpolation. The reduced visual distraction from camera motion and film defects plus the accuracy of our result to the camera original allows for clearer observation of both static and dynamic features of the filmed subject in future analyses.

KEY WORDS: Feature detection, keypoints, homography, signal-to-noise ratio

INTRODUCTION

The Patterson-Gimlin Film (PGF), photographed on October 20, 1967, by Roger Patterson with Bob Gimlin as witness and co-participant, is a historical artifact that needs little introduction. Countless amateur and published analysis efforts to confirm or debunk whether the film portrays a natural biological entity exactly as it appears, or an actor in a costume, have been undertaken over the last half century. Meldrum (2006) devoted a chapter to scientific reaction and critique of the PGF, while offering his own analysis of the comparative functional morphology and kinematics of the film subject. Munns and

Meldrum (2013a) analyzed the integrity of the physical film itself as well as anatomical comparisons between the PGF subject and extant apes and humans in two separate publications. Munns (2014) separately published a non-peer-reviewed book detailing the history of the state of the art in creature suit design and the implications of the PGF subject in the context of the practices at the time.

While the PGF source material and attempts at anatomical and biomechanical analysis are not novel, each succeeding generation of research gains access to tools and techniques originating in other disciplines that allow for new interpretations of the same data. Previous analysis methods in both published and self-promoted works relied on ad hoc manual manipulation to stabilize the camera shake and reduce image defects present in multi-generation film copies. Such methods were seminal in bringing attention to the biological plausibility of the subject depicted in the film but were also labor intensive and prone to errors due to both individual bias and the imprecision of visual approximation compared to algorithmic solutions.

Previous work was also limited by the image quality of the source material. A long-standing constraint of PGF analysis is the absence of the camera original film reel. As opposed to digital media, film loses clarity and gains noise with subsequent copying and is also subject to physical damage, such as scratching and tearing. Even if the camera original were procured, unless it has been carefully preserved in a climate-controlled environment it is likely to be substantially degraded from its original state.

We improve on both these limitations with two contributions:

1. We create composite frames using 8 independent high quality, zoomed-in prints believed to be second generation copies from the ANE group (Munns and Meldrum, 2013a). The copies were digitized by photographing each frame individually at close range with a 12-megapixel (MP) digital single lens reflex (DSLR) camera. Each frame index is aligned to all other copies with a mathematically optimal transformation and merged into a single frame to produce a composited result that combines the image detail from 8 different copies while increasing the signal to noise ratio (SNR) by a factor of $\sqrt{n}$, where n = 8 copies.

2. We stabilize each frame to a reference background scene set as a high-quality first-generation scan of frame 354 (F354). While previous manual alignment attempts only allowed for in-plane rotation and 2D translation for frame alignment, our computational method uses the RANSAC algorithm to solve for the full homography matrix that best aligned detected image feature pairs between individual frames and a stable background scene. This process allows for both pitch and yaw of the frame for a mathematically optimized 3-dimensional alignment.

To our knowledge, individual frame SNR enhancement has never been achieved due to the lack of access to multiple high-quality scans of real film reels and the prohibitive amount of labor involved to accomplish an 8x align and composite for a large frame count. Previous attempts to clean up details in the PGF using image editing software can manipulate the noise and film artifacts present in a single copy but lack the ability to fill in missing data with pieces from other copies or boost SNR through multiple sampling. Computational stabilizations of the PGF have been attempted in recent years, with results posted on YouTube for public viewing. These stabilizations were performed on low resolution digitized single copies of the film made available for previous television productions with unknown copying histories and suffer from the same frame quality issues as the manual analysis efforts.

Furthermore, computational stabilization was sometimes achieved with a convolutional neural network (CNN) and not an explicit modeling of the feature detection and 3D transformation. These architectures invariably introduce warping, distortion, and artifacts as they are globally visually convincing on full frames but locally very unstable on details.

Our method produced an SNR boosted composite frame sequence that was then stabilized as a playback with no artifacts or nonlinear distortions guaranteed by virtue of the mathematical model (Figure 1). We encoded our result to play back at 16 frames per second (fps) at 4k (3840 x 2160) resolution with no false details introduced by AI pixel supersampling or frame interpolation as is the case in prior unpublished attempts. This framerate was chosen based on the inferred camera running speed of the original K-100 camera, which was 16 fps at the minimum and was typical at the time for amateur shoots with no accompanying soundtrack (Munns 2014). Contrary to previous computational stabiliza-tions, our results are suitable for future close analysis of the anatomy and biomechanics of the film subject.

METHODS

Our methodological approach is summarized in Figure 2. Summarized it consisted of two main parts:

1. Cleaning up individual frame quality by reducing noise and filling in missing data with multi-sample frame compositing
2. Producing a high-quality video playback from the resulting frames with homography optimized motion stabilization.

Multi-Sample Frame Compositing

Individual multi-generational copies of the PGF accumulate random noise and non-random film artifacts and lose detail that was present in the camera original with each successive copy. We mitigate the loss of detail in each individual copy by merging multiple high-resolution scans of different copies together. For each frame, we align all copies of that frame to a single master frame and merge them down into a multi-sample composite frame with a sampling factor of 8.

Individual frames benefited from multi-sample compositing in two main ways:

1. Reduction of non-randomly distributed defects, such as scratching or holes in the film
2. Reduction of randomly distributed noise, both from the film copying process and the digital photography

Non-random defects are the result of physical damage to the film and do not occur uniformly within a frame or across different copies. These types of defects simply change the presence or absence of pixel data for a given frame in a given copy, typically by replacing the pixels with near black or near white damage artifacts. These artifacts are rarely present in the same place across multiple copies as they are the result of local contact damage and not global copying loss. A multi-sample compo-sition of a frame will produce a result that has n "votes" for a color value at each pixel location, with the final composite result being the average. Outlier values such as damage artifacts will usually only hold a $1/n$ weighted vote in the average and thus be suppressed by a factor of n with more frame samples, provided the copies were independently made. A visual representation of filling in missing or damaged pixels with information from other copies is shown in Figure 3.

Global, randomly distributed noise (akin to static on an old TV) is the result of copying loss as well as digital photography sensor noise. Film grains are not uniformly arranged in fixed arrays and result in lossy reprojections where these grains misalign when copied from one

generation to the next. Furthermore, all digital devices exhibit some amount of dark current noise, visible as individual fluctuations in pixel color when taking a completely blacked out picture with the lens cap on. This can be probabilistically modeled as Gaussian noise with unknown mean and variance. Since we solve for the color transformation from each copy to the target background, we approximate the noise mean as 0 indicating zero color bias. A copied frame F can then be represented as the original frame μ plus random noise with standard deviation σ.

$$F = N(\mu, \sigma) \quad (1)$$

The variance $\sigma_{\bar{F}}^2$ of $\bar{F}$, the average of n independent samples of F with the same standard deviation σ, is σ^2/n by the definition of the variance. The signal-to-noise ratio (SNR) of $\bar{F}$ can be defined as

$$\mu / \sigma_{\bar{F}} = \mu\sqrt{n} / \sigma \quad (2)$$

which increases by a factor of $\sqrt{n}$ when compared to the single frame representation. Thus, physical defects of the film are suppressed and only have a magnitude of $1/n$ in the composite when present and the SNR resulting from random copying noise should increase by a factor of $\sqrt{n}$, which is 2.83 for n=8. A visualization of reducing random noise through multi-sample compositing is shown in Figure 4. Converging to the signal mean via averages of large numbers of samples is justified by the Law of Large Numbers.

We used the digitized PGF frame archive, collected as described in *When Roger Met Patty* (Munns, 2014), as the source material for our analysis. These scans represent the most widely sampled and highest quality digital scans known to date. Each individual frame was hand spooled across a backlight and photographed with a Canon EOS Digital Rebel XSi DSLR (resolution at 4272 x 2848). The camera parameters were as follows: f/7.1, ISO-1600,

exposure 1/80s, and 100mm focal length. The images were saved as raw .CR2 files as well as JPG images. We worked with the camera-provided JPG encoded images. We also produced alternative versions with PNG encoding from the .CR2 files. This changes the white balance and exposure from the JPG encoding but does not visually change the compression quality given the limited resolution of the film itself. We used the archive copy numbers 8 and 14, known internally as the ANE group. These two copies provided a total of 8 samples of 150 frames before and after the reference lookback frame at F354. Each copy was printed in 4x slow motion for use during a time in which broadcasts were printed on real film stock. Since the reel must run at a consistent speed, a 0.25x slow motion sequence required each frame to be reproduced 4 times in series. Each reprint was a separate sample of the previous generation's frame with independent film grain noise and potential damage. Thus, for our purposes they are each four independent copies for a total of 8 samples per frame. A close-up of a frame from Copy 8 is shown in Figure 5.

For each frame, the samples were aligned to a master copy computationally. Image feature detection was performed on both query and master frames using the Scale Invariant Feature Transform (SIFT) algorithm (Lowe, 2004). SIFT is a computer vision feature detection algorithm that converts image features into numerical vectors whose similarities can be compared across different image detections, known as keypoints. SIFT was performed on all copies of a frame, and feature matching was performed between n-1 copies and a single master copy by finding the nearest neighbors of the keypoints of each frame pair. The choice of master copy is arbitrary: we picked the first frame of Copy 8 in each 4-frame series as the master. We used the OpenCV implementation of SIFT with default parameters and programmed our solution entirely in Python 3.4. A visualization of feature detection and

matching between two copies of the same frame is shown in Figure 6.

We solved for the planar homography between each query/master frame pair (7 total pairs) using the RANSAC algorithm from the matched feature detections in each pair. A 3D homography matrix representing the rotation and translation difference between two planes, the master and the query frame, can be calculated for each pair (Szeliski, 2010) and applied to each query frame to align its detected features to the corresponding features in the master copy. This transformation accounts for the slight differences in frame alignment and rotation relative to the DSLR sensor during scanning and can be calculated for the 1200 total frames in a manner of minutes. The homography matrix is a 3x3 matrix that relates the 3D difference between two planes. A visualization of a homography transformation is shown in Figure 7.

A color correction operation was necessary due to the different biases introduced by the different film stocks and copying methods. This was done so a composite frame composed of many different source copies could be superposed onto a background frame without excessive visual distraction. We approximated the color shift of each frame with a linear equation

$$c' = mc + b \qquad (3)$$

where c' is the target background color, c is the individual copy color, and m and b are the slope and intercept respectively of a linear relationship. For a single pixel index, this can be expressed as:

$$\begin{bmatrix} m_r & 0 & 0 & b_r \\ 0 & m_g & 0 & b_g \\ 0 & 0 & c_b & b_b \end{bmatrix} \begin{bmatrix} c_r \\ c_g \\ c_b \\ 1 \end{bmatrix} = \begin{bmatrix} c_r' \\ c_g' \\ c_b' \end{bmatrix} \qquad (4)$$

We solved for m and b for each of the three channels R, G, B independently by aligning F354 from each copy to the background F354

scan using the feature detection method described in the previous section. We take c and c' from each pair of aligned pixels and create a 3n x 4 matrix by stacking all instances of equation 3 in row major order, where n is the number of aligned pixels in the image pair. We subsampled the pixels as using all pixel correspondences resulted in a matrix with n on the order of 10^6. We solve this system of linear equations with least squares optimization.

Although this color relationship is not guaranteed to be linear, it is a conservative model that can be solved quickly with linear matrix operations. Color mappings involving high order terms such as quadratics can be explored in future work. A demonstration of linear color correction is shown in Figure 8.

Homography Optimized Motion Stabilization

The substantial motion in the original film makes it difficult to observe the subject in frame without excessive visual distraction. Furthermore, digital scanning of the film introduces yet another source of frame motion as the plane of the film is not guaranteed to be perfectly orthogonal and center aligned to the camera image plane in each copy. We correct for this motion using feature detection and homography transformation as described in the previous section.

In this use case, we take each composited frame and detect feature pairs between it and a single master background frame, in this case a high-resolution digital photo of a 1st generation 4x5 inch print of F354 assembled by Bill Munns. We solve for an 8-degree-of-freedom homography matrix for each of the 150 frames that transforms each frame to the same perspective viewing plane as the background scene. This method is more accurate than previous attempts at stabilization using manual alignment due to its ability to account for two extra axes of rotation and perspective distortion, shown in Figure 9. Previous attempts could only rotate each frame in the

plane of the screen (the red Z axis) while neglecting the possibility of pitch and yaw and their associated perspective distortions.

We cropped the stabilized composite frames to a 16:9 ratio in a region zoomed in on the walk cycle in the 150-frame sequence and scaled the resolution to 4K (3840 x 2160). The cropped region was 3500 pixels wide, so a modest upsample of less than 10% using the Lanczos-4 algorithm was necessary.

RESULTS

We merged 8 copies of 150 frames (1200 frames total) into a single stabilized 150 frame composite sequence cut to 4K resolution and encoded it at 16 fps with *ffmpeg*. Although not as clear as the camera original, whose sharpness can be estimated with the clarity of the 4 x 5-inch first generation print used as the background canvas, our result is noticeably clearer than any individual copy in our archive and any copy of the PGF shown to the public in the past. Our results may approximate the appearance of a clean, undamaged 1st generation copy, although this is difficult to confirm without the documentation of a confirmed first-generation reel.

We performed a synthetic benchmark using the 4x5 inch print of frame 354 to test the effectiveness of our method on gaussian noise. We randomly generated 8 lossy copies of this frame by adding to each pixel a random value sampled from a normal distribution $N(0, 25)$. We merged the 8 copies into one composite image and compared the mean squared error (MSE) for pixel difference between the composite image and the original, as well as each of the individual copies. This simulates our method on an artificially generated dataset with a noise-free ground truth. The MSE for all 8 individual noisy copies was 610, while the MSE for the composite was 77. This represents an 87% reduction in MSE in the composite image. The peak signal-to-noise ratios (PSNR) were 19.2 for the noisy copies and 28.2 for the

composite, a 1.47x increase. This is below the theoretical 2.83x increase. Doubling the standard deviation of the noise to 50 does not change the MSE ratio but improves the PSNR ratio to 1.67. Our method may be more effective as the level of grain noise in the image increases, which is a desirable trend. These results are illustrated in Figure 10.

A close-up before and after comparison of the film subject is show in Figure 11. Before frames were sampled from the first frame of Copy 8 in the 4-frame sequence (C8-1). After frames are shown after an 8x alignment and composition with a linear color correction applied. Note the reduction of grain noise and suppression of physical scratches.

Three before and after frames showing full-frame background detail are shown in Figure 12. Note the elimination of the film defects around the right hip of the subject in the second and third rows.

A single frame example of the homography optimized alignment is shown in Figure 13. Keypoints from feature detection are not shown, but the final position of the corners of the query frame in the background frame is indicated. Frames were cropped before stabilization to remove the black film borders from the results. The composite frame on the left demonstrates an extreme example of filling in missing detail with other frames in the black margins, where half of the copies (all 4 prints of C8) excluded that portion of the frame while it was present in the other half.

A full 20 frame walk sequence is shown in Figure 14 to demonstrate the stability of homography optimization. Solving for the transformation matrix between composite frame and background frame produced much closer representations of 3D camera motion than manual manipulation. These results show that our method is robust to heavy distortion from motion blur.

Details observed in these restored frames can be annotated with much higher confidence than in previous work, as features present in our

results are likely to have persisted across multiple independent copies rather than simply being the product of an isolated copying error or damage artifact. Munns and Meldrum (2013b) remarked on the necessity of observing any proposed anatomical feature across multiple independent copies to establish confidence that it is in fact a feature of the true photographic record. Our results fold this cross referencing into a single frame sequence by heavily suppressing artifacts present only in isolated instances while amplifying the image signal that is common to all copies.

DISCUSSION

Advances in computational algorithms and the aggregation of many high-quality digital scans of various film copies allowed us to produce a clearer and more motion stabilized rendition of the PGF than what manual manipulation of singular degraded copies were capable of in the past five decades. Although this effort alone cannot resolve the question of what was really filmed on October 20th, 1967, a mathematically optimized restoration of the film will hopefully allow a higher level of future analysis of and debate about the film subject absent a layer of visual distraction from image noise and motion that was assumed to be inextricable from the image data in the past. Furthermore, our restoration is reliably reproducible due to its foundations as a numerical optimization algorithm and is not subject to individual biases or errors that inevitably arise from various independent attempts to stabilize and enhance the film by hand in photo-editing software.

We did not include additional image manipulation tricks utilizing artificial intelligence (AI) and deep-learning neural networks as part of our main method or results as these techniques create distortions and false details that are not reliable for anatomical or biomechanical analysis. Two common forms of video enhancement with deep learning are pixel interpolation and frame interpolation.

The first stretches the image to a higher resolution by expanding the original pixels to fill the desired dimensions, then making up the pixels in between using learned convolutional filters trained on thousands of other images. This is sometimes referred to as "AI superresolution" and is very susceptible to introducing false details. If an image is doubled in width and height, then a full 75% of all pixels in the result were generated by the deep neural network in an attempt to fill in the blanks between the low-resolution pixels. This contrasts with simply resizing the image with an interpolation algorithm, which uses a defined local mathematical function to transform the image to a new larger or smaller size without losing or introducing detail. We did not use superresolution because our frames were scanned at above 4K resolution, and because the result would not be faithful to the camera original.

The second form of AI enhancement, frame interpolation, generates entirely new frames between existing frames to create a playback that appears smoother than the original sequence. The generated frames in these methods are even less reliable than superresolution techniques because 100% of the resulting pixels in the new frames are synthesized. This can be applied recursively to exponentially increase the framerate of a video sequence, at the cost of reducing the fraction of ground-truth frames in the playback. At 4x interpolation playing at 64 fps, three out of every four frames are synthesized interpolations. (The 2x interpolation is doubled again, but with even less precision as each frame pair contains one real and one generated frame)

We processed our 16-fps video result with an implementation of RIFE-CNN (Huang, 2020) to produce 32 fps and 64 fps high framerate playbacks as an exercise. The results are pleasing to the eye as it reduces the choppiness of the original playback rate

substantially, but we maintain that any future analyses must refer back to the original 16 fps frame sequence to ensure all proposed features were present on the original film reel.

At most, the high framerate versions can help clarify or identify the motion and dynamics of the subject by smoothing out the playback in a way that is less distracting to a viewer's eye. All hypotheses regarding dynamic features of the subject must be confirmed on the original 16 fps playback to ensure the observations are not of artifacts or distortions injected by the neural network. These hypotheses will still be inherently low confidence as starting an analysis with the high framerate render creates a huge potential for bias towards the predilections of the neural network output. An example of an interpolated frame with a subtle motion artifact is shown in Figure 15.

Our restoration and stabilization method were applied only to 8 copies of 150 frames of the film. As there are over 900 frames in the full sequence and over 20 PGF copies in the digital database compiled by Munns, there remains much room for improvement in the quality of the remainder of the PGF reel. However, regardless of how many copies we use in future multi-sample composites of the remainder of the film, it is likely that the results presented in this work are the best-case scenario for quality restoration. The frames of copies 8 and 14 were originally copied with 4x zoom projection to make the bipedal subject take up more of the frame, meaning the grain density on the subject itself is much higher than any of the other copies and the copying loss is minimal. These copies were also in relatively good condition compared to many of the other full reel scans, which suffer from increased prevalence of scratching and tearing. The use of many copies (up to 20 in some frames) may mitigate some of this damage, but the lower pixel density of the bipedal subject in these full frame copies is a fundamental limitation of the source data.

CONCLUSIONS

We present a method for restoring the Patterson-Gimlin Film to its best-to-date quality approximating that of a clean first generation copy by computationally aligning and merging 8 copies of the film down to one composite. We estimated that multi-sample frame compositing reduced the MSE of grain noise by 87% and physical damage artifacts by a factor of 8. We stabilized the composited frames onto a high-resolution background canvas by solving for a homography matrix that produced a mathematically optimal alignment on detected feature pairs between each composite frame and the stationary background. Our method was based on analytical, well understood, highly cited computational methods that are mathematically sound and did not rely on unverifiable deep learning filters to generate new pixel data. We believe this is the clearest, best stabilized, and most accurate version of the PGF rendered to date due to the quality and quantity of our source material and the mathematical optimality of our method. Future work can extend this method to the full 900+ frame PGF sequence or do more in-depth analysis of the proposed anatomical and biomechanical features observed in the film subject.

ACKNOWLEDGEMENTS

LITERATURE CITED

Huang Z, Zhang T, Heng W, Shi B, and Zhou S (2020). RIFE: Real-Time Intermediate Flow Estimation for Video Frame Interpolation. *arXiv* preprint *arXiv*:2011. 06294.

Lowe DG (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision,* 60(2):91-110.

Meldrum J. (2006) *Sasquatch: Legend Meets Science.* New York: Tom Doherty Associates.

Munns B and Meldrum J (2013a). Analysis Integrity of the Patterson-Gimlin Film Image. *The Relict Hominoid Inquiry.* 2:41-80.

Munns B and Meldrum J (2013b). Surface Anatomy and Subcutaneous Adipose Tissue Features in the Analysis of the Patterson-Gimlin Film Hominid. *The Relict Hominoid Inquiry.* 2:1-21.

Munns W (2014). *When Roger Met Patty.* Charleston, SC: Createspace Independent Pub.

Szeliski R (2010). *Computer vision: algorithms and applications.* London: Springer.

Isaac Y. Tian, MS is a doctoral candidate at the University of Washington Paul G. Allen School of Computer Science & Engineering. His research focuses on applications of computer vision, computer graphics, and machine learning for modeling and estimating human body geometry and correlated health metrics from 2D and 3D optical images. He is published in the journals Medical Physics, Obesity, and the Journal of Natural Products for work in computation and AI assisted medical research. He earned his BS in Electrical Engineering and Computer Science at the University of California, Berkeley with an honors concentration in biomedical engineering and a minor in music. Although not formally trained as a biologist, he is an avid enthusiast in paleontology and natural history, has taken coursework in paleoanthropology, and is a dedicated fan of dinosaur research and paleoart.
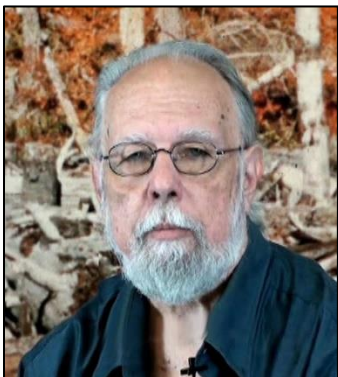
William "Bill" Munns is a veteran motion picture special makeup effects designer with decades of experience fabricating various makeup effects and "creature costumes" for motion pictures, television and commercials, as well as designing robotics for theme parks and museums. To analyze the subject figure seen in the Patterson-Gimlin Film, he has applied his extensive knowledge of the processes, materials and techniques of creating ape-like fur costumes for human performers to wear. And as both a vintage filmmaker (knowledgeable in making and editing 16mm films) and a computer graphics artist, he possesses unique skills and knowledge vital to analyzing the Patterson-Gimlin Film itself, not just the subject figure seen within it. He has spent 14 years analyzing and researching the film, and thus brings a wealth of knowledge and experience to the subject.

Jeff Meldrum is a Full Professor of Anatomy & Anthropology at Idaho State University (since 1993). His research centers on primate locomotion generally, and the evolution of hominin bipedalism specifically. His professional interest in sasquatch began when he personally examined a line of 15-inch tracks in the Blue Mountains of southeastern Washington, in 1996. Over 25 years later, his lab houses well over 300 footprint casts attributed to relict hominoids around the world. He conducts collaborative laboratory and field research throughout the world, and has shared his findings in numerous popular and professional publications and presentations, interviews, and television appearances. He is author of *Sasquatch: Legend Meets Science* (Tom Doherty Publishers, 2006) and the editor-in-chief *The Relict Hominoid Inquiry* (www.isu.edu/rhi).

Figure 1. Comparisons of a degraded 3rd gen copy (**A**) to our result (**B**) and a 1st generation, almost lossless 4x5 inch print (**C**). Note the reduction in scratches, appearance of static noise, and correction of washout colors.
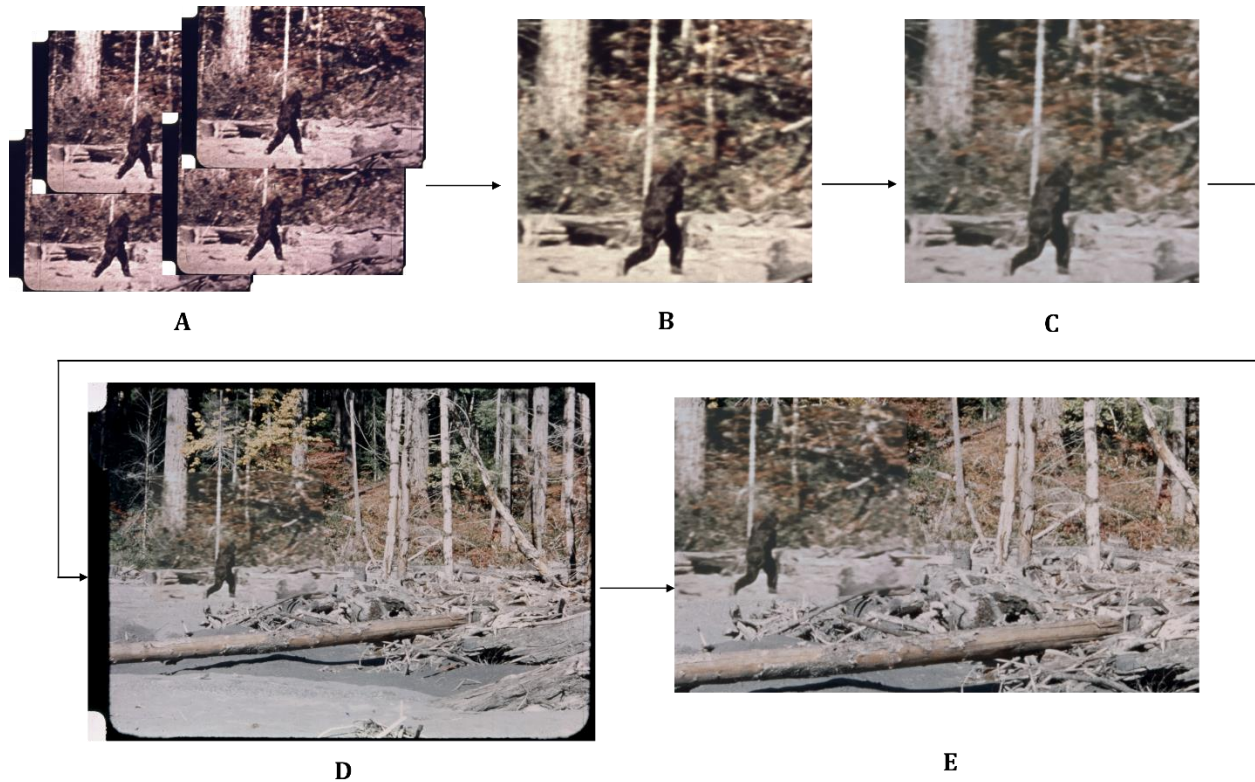
Figure 2. Visual overview of method. **A**. Digital photos were taken of multiple physical copies of the same frame of film. Each copy contains different grain noise and film damage. **B**. Copies are computationally aligned and merged down into one frame using SIFT + RANSAC algorithms. This amplifies the film details that are common to all copies while suppressing the noise that is only present in a single copy. 8 copies were used in this work. **C**. We solved for a linear color correcting matrix for the red, green, and blue channels independently to correct the composite frame's appearance to the target background frame, which is a high-resolution scan of a first generation 4x5 inch print of Frame 354. **D**. We aligned the color-corrected frame to the background frame with SIFT + RANSAC, which finds the best alignment based on the background scenery common to both images. **E**. We crop a 16:9 aspect ratio portion of the frame, resize it to 4K resolution, and compile the 150 total frames into a video playing at 16fps.
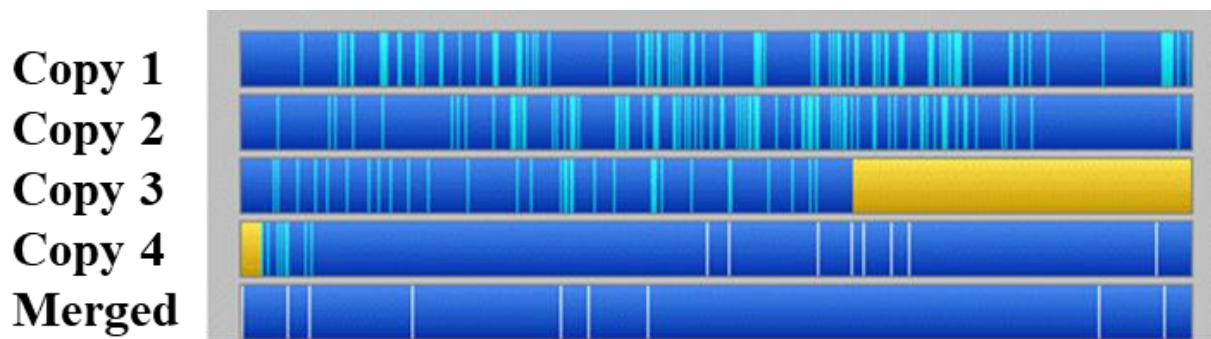
Figure 3. Restoration of non-random defects by filling in damage (light blue) or missing data (yellow) with the sum of image data from all copies.
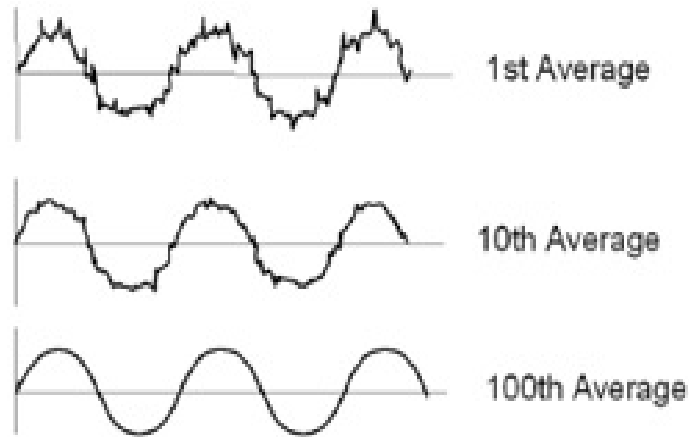
Figure 4. Average of n samples boost SNR by a factor of $\sqrt{n}$. This is analogous to the signal processing equivalent known as time synchronous averaging, where a noisy signal is averaged with repeated samples of itself to cancel out perturbations. The convergence of the average of n samples to the true mean value as n grows towards infinity is consistent with the fundamental principle of probability known as the Law of Large Numbers.

Image credit: https://www.crystalinstruments.com/time-synchronous-average
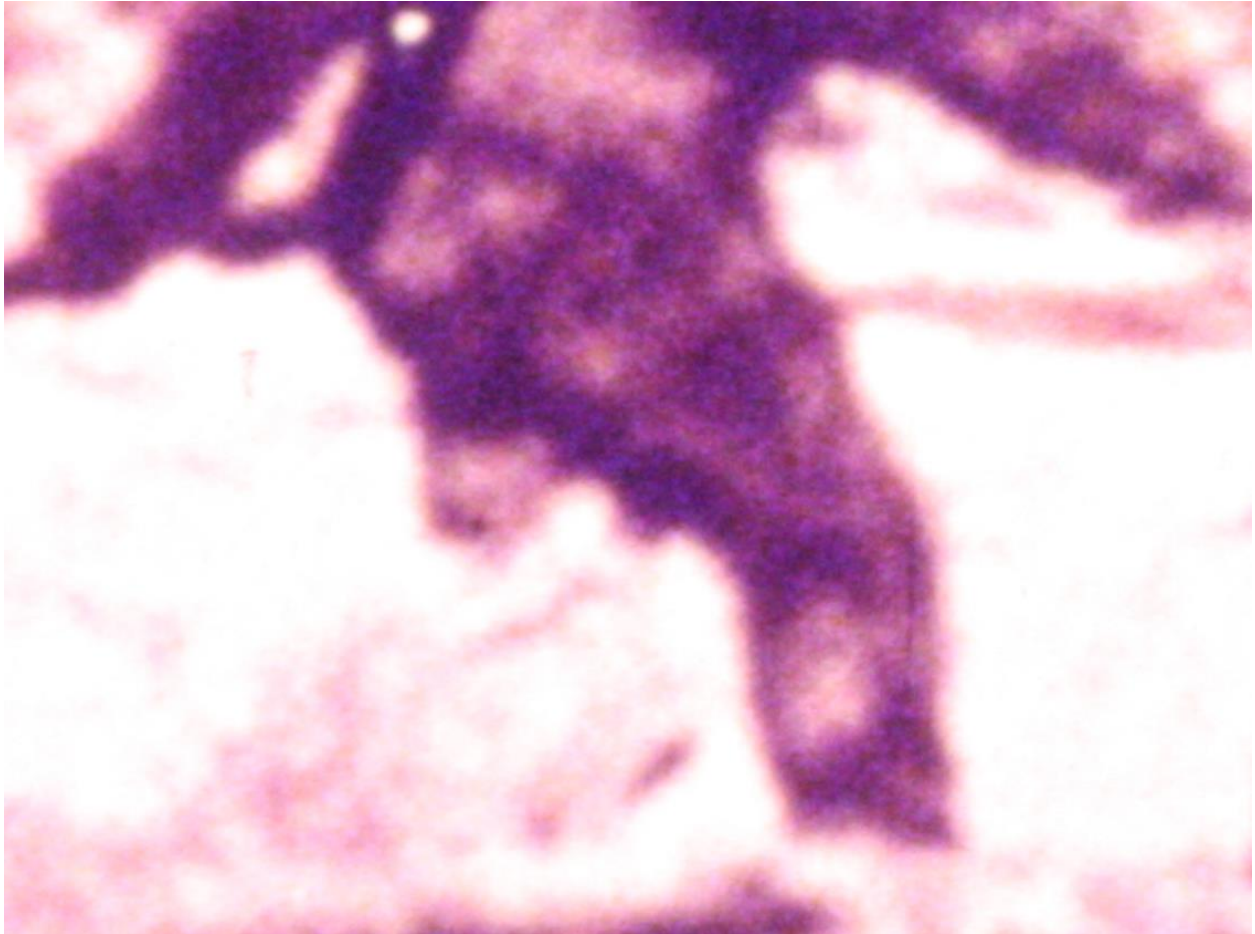
Figure 5. Zoomed in frame from Copy 8 showing pixelated random noise and a white pockmark from film damage near the right arm.
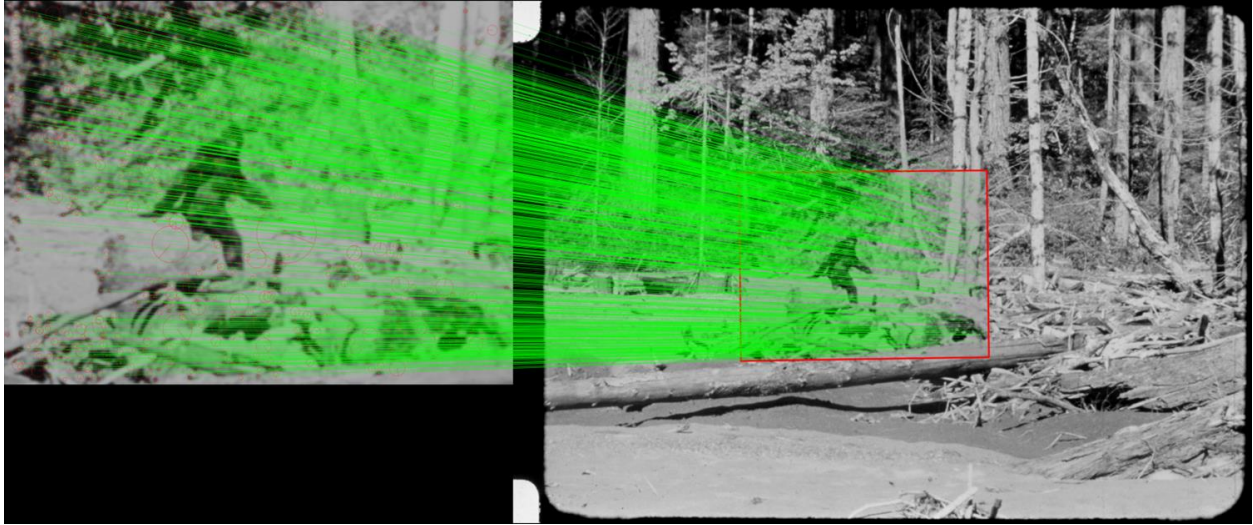
Figure 6. SIFT feature detection example between two copies. Red circles indicate keypoint with orientation, green lines indicate nearest neighbor matches.
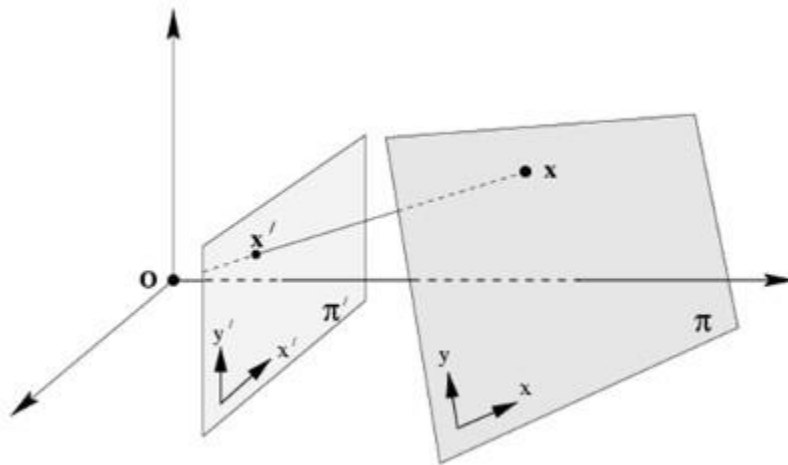
Figure 7. Diagram from the OpenCV documentation showing how homography matrix H relates keypoint **x** in the query frame to the matched keypoint **x'** in the master frame.

Figure 8. Left, a frame from Copy 8 with an elevated pink hue. Right, after color correction to the F354 background.
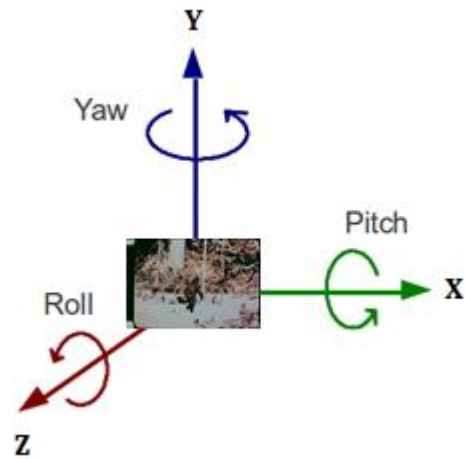
Figure 9. Manual on-screen manipulation in photo editing software only allows for rotation in the Z-axis (roll), whereas a 3D transformation contains pitch and yaw.

Figure 10. Left, the original 4x5 print. Middle, a noisy copy generated by adding $N(0, 50)$ to all pixels. Right, result of merging 8 independent noisy copies into one frame. MSE reduced by 87% and PSNR increased by a factor of 1.67 relative to the single noisy copy.

Figure 11. Before and after multi-sample composition. Note the vertical copying defect and diagonal blemish on the original copy.
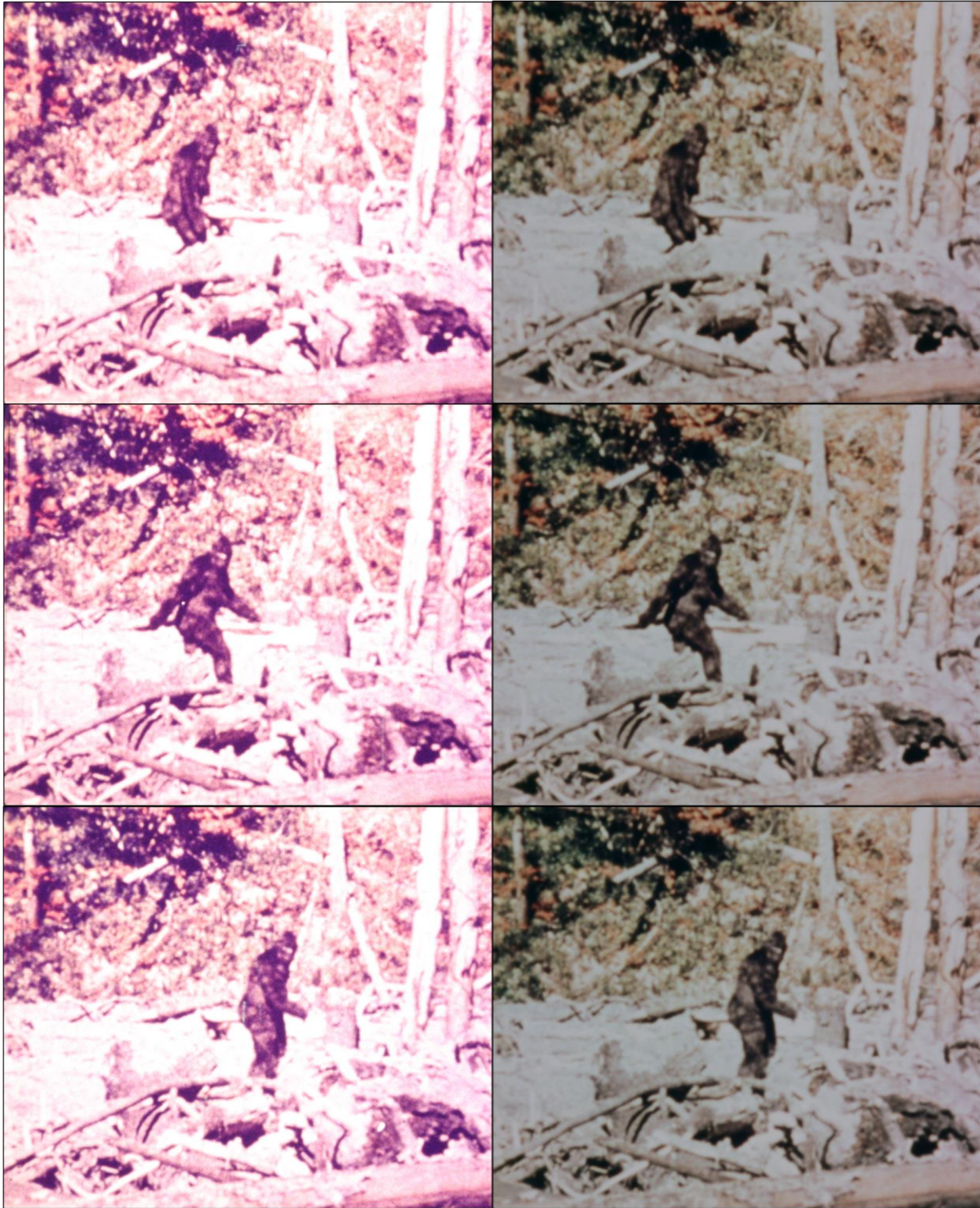
Figure 12. Six frames showing the before (left, single master copy) and after (right, 8x composite) frames after alignment, composition, and color correction. Note the reduction in the appearance of static noise and physical damage.
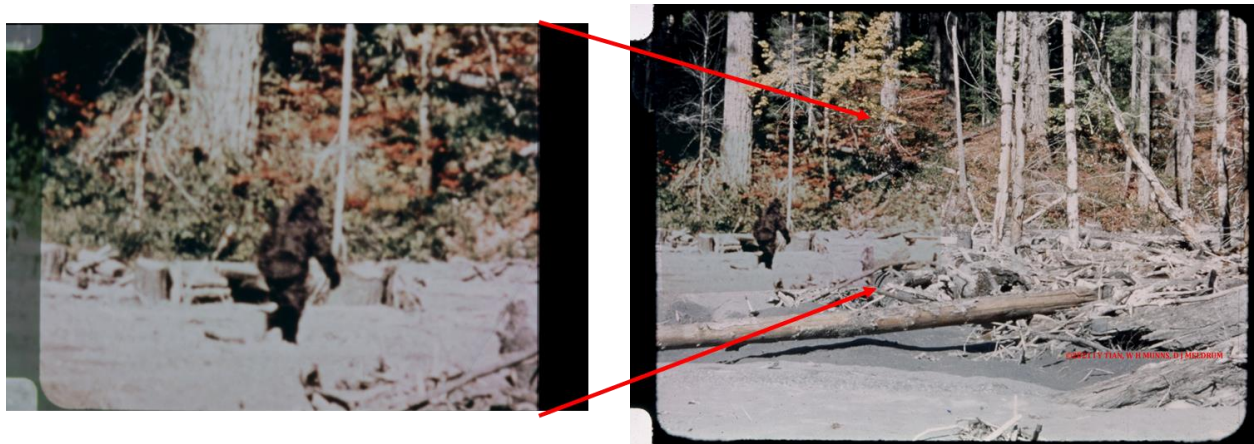
Figure 13. Motion stabilization of a single frame using homography. Red arrows show where the corners of the query frame were pasted into the background frame.
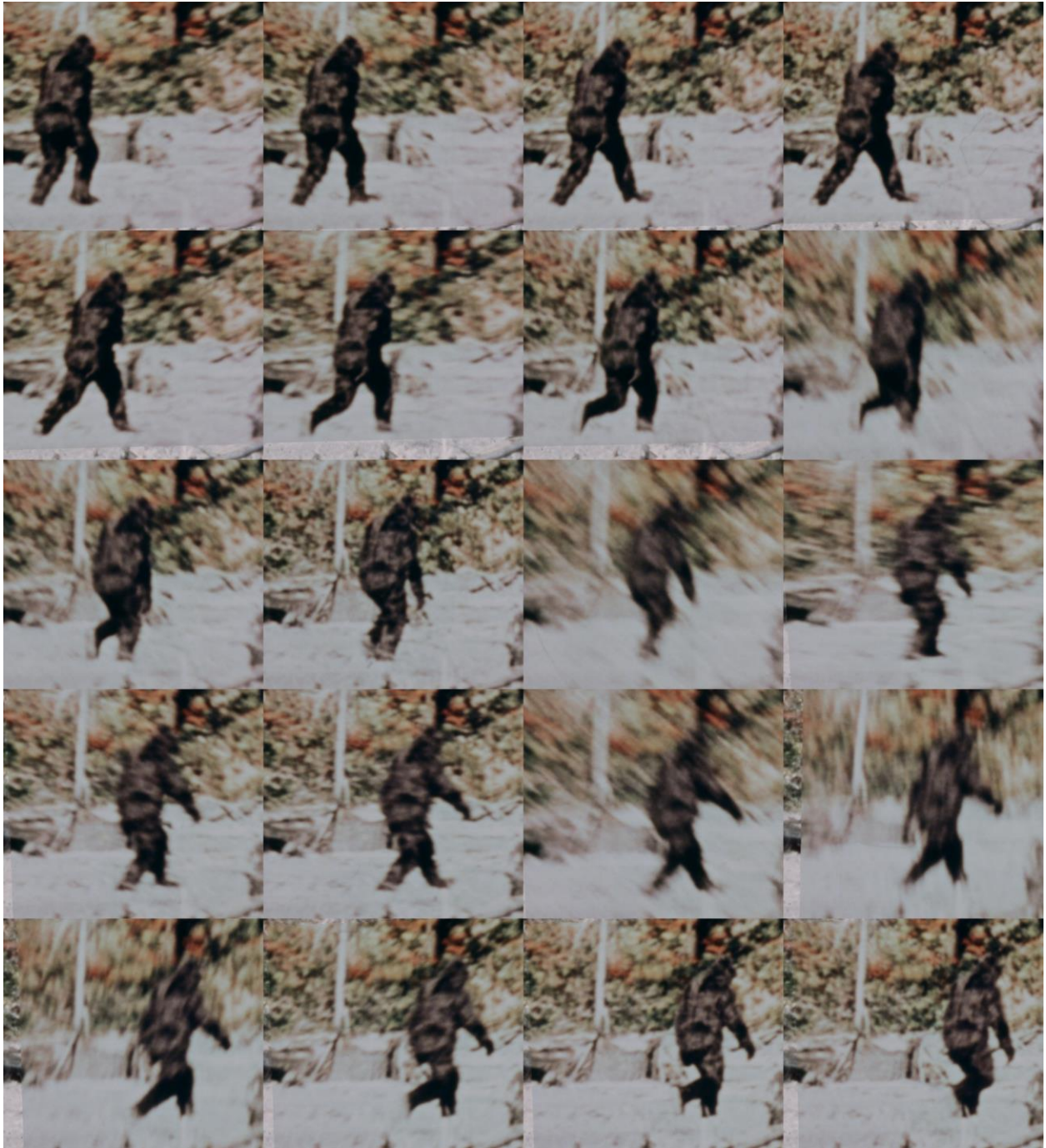
Figure 14. Twenty consecutive frames demonstrating the effects of homography computed motion stabilization. Notice how even in instances of extreme motion blur the walk cycle is relatively stable compared to adjacent frames.

Figure 15. An interpolated frame from a 2x interpolated, 32fps RIFE-CNN processing of our result. Although most of the body is visually indistinguishable from an original frame, there is an artifact introduced by incorrect motion tracking and interpolation of the arm circled in red. Other smaller distortions may be too subtle to see and cannot be depended on for ground truth information.