Image averaging and shape recognition in forensic analysis of video evidence

Geoffery T. Desmoulin* and Mark Nolette

GTD Scientific, Inc.,
2037 Mackay Avenue,
North Vancouver, BC, V7P 2M8, Canada
Email: gtdesmoulin@gtdscientific.com
Email: manolette@gtdscientific.com
*Corresponding author

Theodore E. Milner

GTD Scientific, Inc.,
2037 Mackay Avenue,
North Vancouver, BC, V7P 2M8, Canada

and

Department of Kinesiology and Physical Education,
McGill University,
Montreal, H2W 1S4, Canada
Email: temilner@gtdscientific.com

Abstract: Averaging of raw images obtained from single frames of a body worn camera was used to illustrate the ability of averaging to significantly reduce the noise present in a single video frame. The outline the subject in the averaged image was then compared to outlines of a different subject adopting four similar postures, which included one posture that mimicked the posture of the first subject (reference posture). Using Hausdorff distance as a measure of shape similarity, it was possible to demonstrate that the outline of the second subject was most similar to the outline of the reference posture when the second subject mimicked the reference posture. Despite the similarity of the four postures, the Hausdorff distance was at least 50% less when the second subject mimicked the reference posture than for any of the other postures.

Keywords: image averaging; image shape recognition; video analysis; shape difference measure; posture analysis.

Reference to this paper should be made as follows: Desmoulin, G.T., Nolette, M. and Milner, T.E. (xxxx) 'Image averaging and shape recognition in forensic analysis of video evidence', *Int. J. Forensic Engineering*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Geoffery T. Desmoulin is the Principal of GTD Scientific Inc. GTD offers Biomechanical Consulting Services throughout North America. Practice areas include Injury Biomechanics, Incident Reconstruction and Physical Testing with a sub-specialty in the Science of Violence®. GTD has been retained in significant complex injury litigation cases involving police department use of force, violent encounters and TASER International to name a few examples. Furthermore, landmark shooting reconstruction methodology developed by Dr. Desmoulin was upheld as reliable and admissible by the U.S. Federal District Court for the 9th District of California. This methodology is published in the *International Journal of Forensic Engineering*.

Mark Nolette, MSc, is a Senior Injury Biomechanist at GTD Scientific Inc. He received a Master's degree in Biomedical Physiology from Simon Fraser University in 2012 and a Bachelor's degree in Mechanical Engineering from Universté de Sherbrooke in 2009.

Theodore E. Milner conducted post-Doctoral research in Brain and Cognitive Sciences and Biomechanics at MIT from 1984-86 and has held positions as Assistant Professor in Biomedical Engineering at the University of Montreal from 1986–1992, Professor in Kinesiology at Simon Fraser University from 1992–2007 and Professor and Department Chair in Kinesiology and Physical Education at McGill University from 2008–2020. He was a visiting researcher in Japan at the ATR from 2000–2007, visiting professor and in Switzerland at EPFL in 2005 and scientific staff member at ETH Zurich from 2014–2016. He is currently senior scientist at GTD Scientific.

1 Introduction

Current forensic investigations often involve analysis of video evidence, which may have limited resolution and clarity. Although it is not possible to change image resolution, since it is a fixed property of the camera from which the images were obtained, there are techniques to improve image clarity (Jähne, 2005). Some techniques rely on image processing which alters the original image by applying various filtering algorithms. Furthermore, the ability to improve image clarity may be limited by the image resolution, i.e., the number of pixels depicting the object of interest in the image. As well as the limitation of image resolution, noise produced by fluctuations in lighting and in the intensity of light reflected from objects in the image, reduces image clarity. Night sky photographers encounter the same issues of image noise and often employ a straightforward technique to improve image clarity which does not alter the pixels of the original images. The technique simply involves acquiring multiple images of the same scene, aligning the images and averaging https://www.cambridgeincolour.com/tutorials/image-averaging-noise.htm.

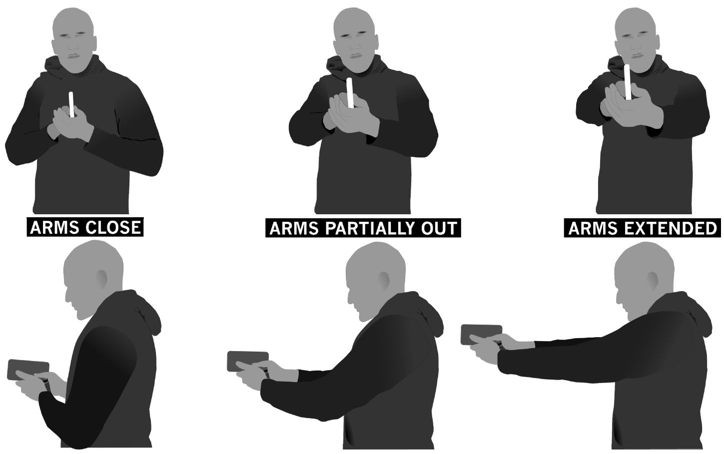
If we assume that the intensity of light reflected from a given point on an image fluctuates randomly then the value of the camera pixel corresponding to that point on the image can be considered as a random variable. The noise associated due to random fluctuations in the intensity of a pixel in a camera image is reduced by the square root of the number of images included in the average (Oliver, 1988, p.165). Therefore, if we have several video frames in which the object of interest is stationary, even if the camera moves slightly, we can align the object of interest in each frame and average the resulting images. The net effect will be to reduce noise in the image and improve image clarity. However, the image average may still lack sufficient resolution or clarity to achieve the desired aim of the forensic investigation. For example, the identity of an object in the image or the configuration of a group of objects may still be unclear. We propose a technique based on image shape outline which can be used to compare the outline of the image under investigation to the outlines of standard images in order to determine quantitatively which of the standard outlines most closely resembles the outline of the image under investigation.

Shape similarity is used extensively in artificial vision and facial recognition. Veltkamp and Latecki (2006) compared the properties and performance of 15 shape similarity measures. Of these, six measures demonstrated uniqueness and robustness to deformation and noise. Of the six, one of the most straightforward to implement is the Hausdorff distance (Huttenlocher et al., 1993; Enayatifar and Abdul Salam, 2013). Although its performance was in the midrange of the other shape similarity measures, we chose to use it because of the simplicity of implementation. We demonstrate its utility for forensic investigation by analysing low-quality video of a scenario in which the Hausdorff distance is used to identify a subject's most likely posture, which has significant implications for the subject's intended action.

2 Materials and methods

To illustrate the effect of shape comparison, we selected a conventional standing posture in which a subject held a mobile phone in four different positions. The purpose was to investigate whether the subject was using the phone like a firearm, to imitate a shooting stance or simply looking down at the phone. Three possible shooting stances, which differ in the amount of arm extension were used in the investigation, as shown in Figure 1. The subject was photographed with an AXON Body 2 camera, while holding a mobile phone in each of the three shooting stances, as well as while looking down at the phone and reading a text message.

Figure 1 Three shooting stances, with varying amounts of arm extension while holding a mobile phone



A second subject, of similar build to the first subject, was later recorded in a short video, using the same camera but under very low lighting conditions in order to illustrate the power of image averaging. The subject adopted the shooting stance shown in the centre panel of Figure 1, i.e., arms partially out. Image averaging was performed for 10, 20 and 30 consecutive video frames, which were acquired at 33 Hz.

The outlines of both subjects in their respective stances, from head to waist, were acquired using the freeform tracing tool after the images had been imported into Powerpoint. The outlines were saved as a.JPG files which were subsequently imported into Inkscape, where they were scaled to the same vertical dimension and digitised using Inkscape's Trace Bitmap and Add Nodes tools. The Add Nodes options were set to 200 segments with a maximum segment length of 10 pixels. The resulting coordinates of the bitmap nodes were saved in a text file.

Shape comparison was carried out by computing the Hausdorff distance (Huttenlocher et al., 1993). The Hausdorff distance between two shapes was calculated with a custom Matlab script which imported the text files of the outline coordinates. One shape, the outline of the second subject, served as the reference. The outlines of the first subject, in the four stances, shown in Figure 2, served as test outlines which were compared to the reference. Each test outline was first shifted horizontally relative to the reference outline. The amount of the horizontal shift was varied from 0 to 50 increments to the left (negative) and from 0 to 50 increments to the right (positive). To calculate the Hausdorff distance for a given horizontal shift, the distance between each coordinate of the reference outline and every coordinate of the test outline was calculated. For each coordinate of the reference outline, the minimum of the distances to the test outline was recorded, resulting in a set of minimum distances to the test outline. The Hausdorff distance is defined as the maximum of this set of minimum distances. The Hausdorff distance was calculated for each horizontal shift increment. The lowest value of the Hausdorff distance represented the best alignment of the test outline with the reference outline. The lowest value of the Hausdorff distance was used as the measure of difference in shape between the reference outline and the test outline. The test outline for which the difference in shape was least was identified as the most likely of the four stances adopted by the second subject.

Figure 2 (A) In the first three panels (left to right), the subject is shown holding a mobile phone in three shooting stances, corresponding to those illustrated in Figure 1 and shown looking down at the phone in the fourth panel. (B) The outlines of the subject are shown below each panel. The darker outline of the extreme left is that of the comparison subject, adopting the shooting stance shown in the centre panel of Figure 1

3 Results

The first subject is shown adopting each of the four different stances in Figure 2(A). The subject was photographed under dim lighting conditions to mimic the ambient lighting that might be present with a body worn camera during a nighttime scenario. The first three panels, from left to right, correspond to the three shooting stances shown in Figure 1. In the fourth panel, the subject is looking down at the phone.

Images from a video recording of the second subject, adopting the shooting stance illustrated in the middle panel of Figure 1, are shown in Figure 3. The video was recorded under very dim lighting conditions. The images in Figure 3(A) represent a single frame and the average of 10, 20 and 30 consecutive frames. Note, that 30 consecutive frames represents 1.0 s of recorded video. The noise evident in a single frame, creates considerable uncertainty as to the subject's outline. In fact, it is unclear whether the image even represents a person. However, the average of 10 images taken over 0.33 s, already produces sufficient clarity to discern the subject's outline. Averaging more frames brings further clarity to details such the identity of the object being held in the hands and facial features.

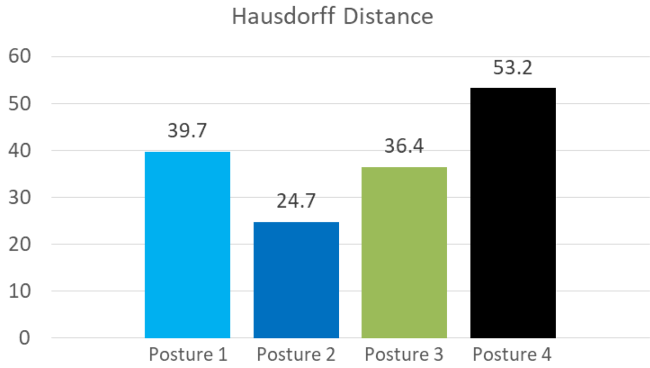
Figure 3 (A) Image of the second subject, adopting the shooting stance of the middle panel of Figure 1, arms partially out, while holding a mobile phone. From left to right: single video frame, averages of 10, 20 and 30 consecutive video frames. (B) Outline drawn around subject which served as the reference outline for calculating the Hausdorff distance (see online version for colours)

A

B

The results of the similarity analysis using the Hausdorff distance are shown in Figure 4. The lowest value of the Hausdorff distance was found between the reference outline and Posture 2 (shooting stance with arms partially out). The highest value was for Posture 4 in which the subject was looking down at the phone. Thus the difference in shape between the reference outline and test outlines was least for the shooting stance with the arms partially out.

Figure 4 Comparison of the Hausdorff distance between the reference outline and the outlines for the four postures shown in Figure 2 (see online version for colours)



4 Discussion

The proposed methodology of frame averaging and shape comparison was implemented in a test scenario, modelling a forensic investigation into an individual's probable posture, relying on video evidence acquired under poor lighting conditions. The methodology was able to clearly identify a reference stance adopted by one subject from among four possible stances adopted by another subject. The Hausdorff distance for the next closest shape, represented a difference of approximately 50% relative to the Hausdorff distance for the closest shape (36.4 compared to 24.7).

The power of frame averaging to reduce noise in an image is clearly demonstrated by the difference in image clarity between the image obtained with a single frame and the image obtained by averaging 10 consecutive frames. Given that body worn camera video is frequently obtained under poor lighting conditions, improving image clarity is a significant concern for forensic investigation. A positive feature of image averaging is that is does not involve any manipulation of the original images. It simply removes random noise in the image. This may be particularly important for judging the admissibility of evidence presented in court.

Although there are a large number of measures which can be used to determine shape similarity (Veltkamp and Latecki, 2006), the Hausdorff distance is one of the most straightforward to implement. In our implementation we used readily available application software, which required coding only to compute distances between coordinate points. This makes it relatively simple to explain the methodology, which can be important for presenting a case. The principal limitation of using the Hausdorff distance in shape comparison is that it is not robust with respect to noise or outliers, e.g., a single outlier can result in a large value of the Hausdorff distance even though the two shapes may be very similar. However, the Haudorff distance computation can be modified to reduce this effect. In particular, computing the Hausdorff distance in both directions, i.e., from image A to image B and from image B to image A and then using the maximum of the two computations will reduce the effect of outliers. In addition, a test parameter such as the average position of the points in image A can be used in computing the distance to image B, if it suspected that there is noise in image A. This will further reduce the effect of noise or outliers.

The ability of the Hausdorff distance to distinguish between shapes could be improved by incorporating an algorithm which learns to recognise a particular shape from among instances of that shape compared to similar shapes. This would involve generating a set of instances of the target shape and similar shapes. In the context of the example presented in this paper, it could be achieved by having a group of subjects of different body builds adopt the target posture and the postures to which the target posture is being compared. A algorithm such as PAC (probability approximately correct) learning could then be used to learn to distinguish the target shape from among the comparison shapes, since PAC has previously been successfully applied to similar problems in computer vision (Felzenszwalb, 2001).

5 Conclusions

Tools to determine shape similarity are particularly useful in forensic investigations where the outline of an object can be traced even if detailed features of the object cannot be clearly identified. The outline can then be compared to outlines of similar objects to identify which shape it most closely matches. The Hausdorff distance is a straightforward, effective measure of the shape similarity that can be applied to forensic investigation.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Enayatifar, R. and Abdul Salam, R. (2013) 'Similarity measure using Hausdorff distance in 2D shape recognition system', *Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation*, Singapore, pp.194–197, https://doi.org/10.2991/3ca-13.2013.49

Felzenszwalb, P. (2001) 'Learning models for object recognition', *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Kauai, HI, USA, pp.I-1056–1062.

Huttenlocher, D.P., Klanderman, G.A. and Rucklidge, W.J. (1993) 'Comparing images using the Hausdorff distance', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, pp.550–863.

Jähne, B. (2005) 'Chapter 10 pixel processing', *Digital Image Processin*g 6th ed., Springer, pp.268–269.

Oliver, W.R. (1988) 'Image processing in forensic pathology', *Clinics in Laboratory Medicine*, Vol. 18, pp.151–180.

Veltkamp, R.C. and Latecki, L.J. (2006) 'Properties and performance of shape similarity measures', in Batagelj, V., Bock, H.H., Ferligoj, A. and Žiberna, A. (Eds.): *Data Science and Classification. Studies in Classification, Data Analysis and Knowledge Organization*, Springer, Berlin, Heidelberg, https://doi.org/10.1007/3-540-34416-0_6