

# Video Recordings of Male Face and Neck Movements for Facial Recognition and Other Purposes

Collin Gros <sup>1</sup> and Jeremy Straub <sup>2,\*</sup>

<sup>1</sup> Department of Computer Science, Texas Tech University, Lubbock, TX 79409, USA

<sup>2</sup> Department of Computer Science, North Dakota State University, Fargo, ND 58108, USA

\* Correspondence: jeremy.straub@ndsu.edu; Tel.: +1-701-231-8196

Received: 21 July 2019; Accepted: 3 September 2019; Published: 6 September 2019



**Abstract:** Facial recognition is made more difficult by unusual facial positions and movement. However, for many applications, the ability to accurately recognize moving subjects with movement-distorted facial features is required. This dataset includes videos of multiple subjects, taken under multiple lighting brightness and temperature conditions, which can be used to train and evaluate the performance of facial recognition systems.

**Dataset:** <https://doi.org/10.17632/xgg8xcscr5.1>; <https://doi.org/10.17632/f47pm7rwt3.1>

**Dataset License:** CC-BY

**Keywords:** facial recognition; males; video recordings; face and neck movement; human recognition

## 1. Summary

This dataset is comprised of video of multiple male subjects' head and shoulder areas while the subjects make regular movements, which can prospectively be used to train and test facial recognition algorithms. The subjects' movements are synchronized across the dataset to facilitate recognition time-based comparisons.

For each subject, data was collected for multiple head positions, light brightness and light temperature. The set of videos includes approximately 33,000 frames, recorded at a rate of 29.97 frames per second. Because these are not discrete images, but instead frames in a video, face detection, and identification during movement and recognition from video can also be assessed using this data.

To this end, the subjects' movements are synchronized between videos to facilitate comparison of algorithms based on total time subject is detected during the video. Specifically, the subjects move their head and neck through an entire range of motion. This facilitates the testing of recognition at different points using individual frames as well as the testing of algorithm performance on moving video data. This dataset can be used with large training sets otherwise available to test recognition with a large database of potential subjects to match to. The data was collected in a controlled environment with a consistent white background.

## 2. Background

Facial recognition has multiple uses including retail store [1,2], access control [3] and law enforcement [4] applications. Facial data can also help classify subjects by gender [5] and age [6] and provide insight into their current interest level and emotional state [7].

While much facial recognition research focuses on static images, many applications require the real-time or near-real-time identification of moving faces from video. Even if single video frames are used and presented to static facial recognition systems, they have movement blur and facial orientations

not supported by the static recognition system. For this reason, facial recognition from video, in many cases with static image training, is a key area of research.

Approaches which use support vector machines [8], tree-augmented naive-Bayes classifiers [9], AdaBoost [10], linear discriminant analysis [10], independent component analysis [11], Fisher linear discriminant analysis [11], sparse network of Winnows classifiers [12], k-nearest neighbor classifiers [12] and multiple other techniques [13] has been proposed. Some techniques have also been developed which take advantage of video properties [14] and motion history [12].

This dataset provides training and testing data for evaluating the performance of video facial recognition systems. In particular, it provides data for both static image and video-based training and video-based recognition. Subjects move their head to multiple positions, facilitating the comparison of algorithm performance with regards to subject head position. Additionally, data collected under different lighting brightness and temperature levels is included to facilitate the evaluation of lighting (as part of training or testing data sets) on algorithms.

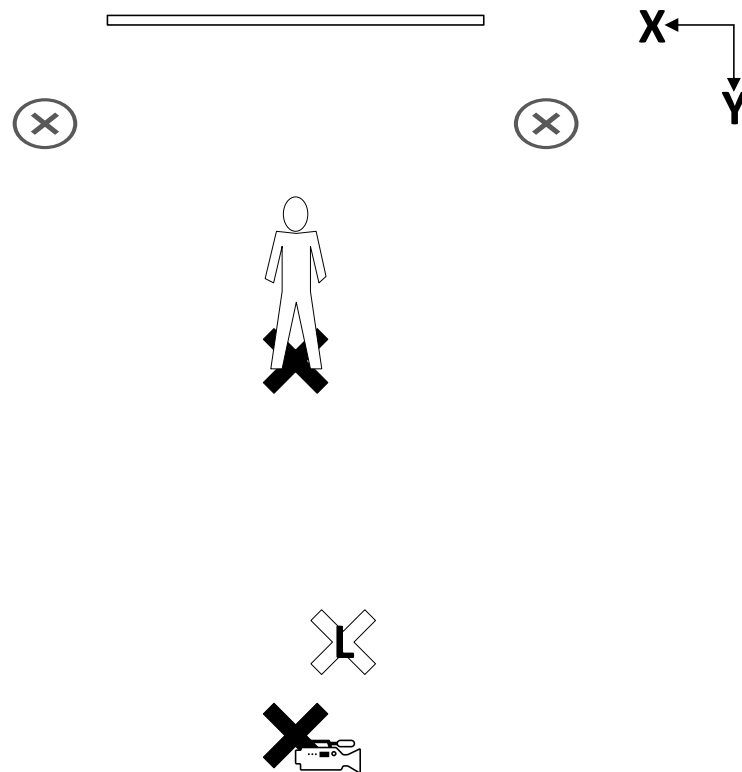
### 3. Experimental Design, Materials, and Methods

This section discusses the equipment, configuration and experimental methods used to collect the dataset. First, the equipment and its configuration are discussed. Then the lighting conditions are presented. Finally, the experimental design is reviewed.

#### 3.1. Equipment and Setup

A Sony AX100 4 K Expert Handycam (Tokyo, Japan) was used for video recording. The high definition 4 K videos ( $3840 \times 2160$  resolution at 29.97 frames per second) recorded were saved in the MP4 file format with the AVC codec. Two Neewer LED500LRC LED lights (Edison, NJ, USA) were used as background lighting. One Yongnuo YN600L LED light (Hot Springs, AK, USA) was used for the lighting of the subject. All lights and the camera were placed at stationary positions, depicted in Figure 1 with location measurements presented in Table 1. A Tekpower lumen meter (Montclair, CA, USA) was used to measure the lumens produced in each lighting configuration. A standard projector screen served as the backdrop for the photos. The audio present in the videos is the sound of an electronic metronome that indicates to the subject when it is time to change positions. This audio was recorded in stereo with a 48 kHz sampling rate using PCM encoding.

One lighting angle was used and the light was pointed directly towards the subject. The camera was placed in a fixed position near the light, also aimed at the subject.



**Figure 1.** Depicts the positions of lights and the video camera.

**Table 1.** Camera, lights and subject positions.

Location	X Coordinate	Y Coordinate
Subject Light	84.5	127.5
Background 1	43.5	50.5
Background 2	129	47
Camcorder	97	132
Subject	96.5	63.5

### 3.2. Lighting

The subject was filmed under multiple lighting levels, which are summarized in Table 2. The lighting of the subject from the Yongnuo YN600L LED light was set at 60% brightness on warm (3200 k), 60% brightness on cold (5500 k), 10% brightness on warm (3200 k) with 10% brightness on cold (5500 k), 40% brightness on warm (3200 k) with 40% brightness on cold (5500 k), and 70% brightness on warm (3200 k) with 70% brightness on cold (5500 k). Lumen readings were measured using a Tekpower lumen meter and these values are also included in Table 2.

**Table 2.** Light setting equipment configuration and measurements of lumens produced for each light setting.

Configuration	Light Settings	Lumens
Warm	60% brightness on warm (3200 k)	280
Cold	60% brightness on cold (5500 k)	391
Low	10% brightness on warm (3200 k) and 10% brightness on cold (5500 k)	155
Medium	40% brightness on warm (3200 k) and 40% on brightness on cold (5500 k)	492
High	70% brightness on warm (3200 k) and 70% brightness on cold (5500 k)	745

### 3.3. Subjects & Procedure

Five videos were taken for each of 11 subjects, using a protocol approved by the NDSU institutional review board, for a total of 55 videos. The subjects were males between the ages of 18 and 26. A few of the subjects had small beards, while the rest have minimal facial hair.

Each of the five videos, for each subject, uses a different lighting configuration. These lighting configurations are depicted in Figure 2. The videos are approximately 20 s long and the subject moves his head to a new standardized position every second. Markers were placed on the walls, floor, and ceilings that the subjects were instructed to look at, to correctly position their heads.



**Figure 2.** The different lighting settings used for each video (left to right: Warm, Cold, Low, Medium, and High).

The subject was told to move their head after immediately upon hearing a metronome-like clicking sound. The time between each tick was one second. Because of this, each subject spends approximately the same amount of time facing in each position and moving between positions. This allows the impact of the lighting conditions on different facial recognition algorithms to be compared in terms of aggregate face detection time, as each view of the subject's face was visible for the approximately the same amount of time in each video. Figure 3 depicts the various positions each subject positioned his head in.



**Figure 3.** Subjects were told to position their head in multiple orientations, for one second at a time, during video recording.

## 4. Comparison to Other Data Sets

A number of data sets exist that have been collected for performing facial recognition work. The majority of these data sets are collections of individual images. Commonly used data sets range in size from the "ORL Database of Faces" which has 400 images for 10 individuals [15] to the "MS-Celeb-1M" dataset which has 10,000,000 images for 100,000 people [16]. Data sets also vary in the number of images that are presented for each subject. The "Labeled Faces in the Wild" data set, for example, included only two images per individual [17]. The "Pgu-Face" dataset had 4 images per individual [18] and the "FRGC 1.0.4" dataset had approximately 5 images per person [19]. Other data sets, such as the "IARPA Janus Benchmark A" [20] and "Extended Cohn-Kanade Dataset (CK+)" [21] have a larger number, including 11.4 and 23 images per individual, respectively. Prior work included producing data sets with 525 [22] and 735 [23] images per subject.

The dataset described herein is comprised of 55 videos (of 11 subjects) which are each approximately 20 seconds in length. Recorded at 29.97 frames per second, this means that each video is approximately 600 frames and that there are approximately 3,000 frames of each subject and 33,000 frames, in total,

for all 11 subjects. This places this dataset towards the larger end of the spectrum, in terms of the total number of frames or images.

Data sets have been collected in a variety of ways. Learned-Miller, et al. [17] created a dataset entitled “Labeled Faces in the Wild” which included 13,000 images for 5,749 people. This data set was harvested from websites. Guo, et al. [16] also created a harvested dataset, “MS-Celeb-1M” of 10,000,000 images covering 100,000 people. Many other datasets are manually collected using volunteer subjects. These databases are typically smaller with images of fewer individuals. The “ORL Database of Faces” contained images of 10 people [15]. Larger datasets included the Georgia Tech Face Database with images of 50 people [24] and the AR Face Database [25] with images of 126 individuals.

Most manually collected datasets present multiple views of the subject, in many cases from different angles. In some cases, lighting or other environmental conditions are varied. The “Extended Yale Face Database B” [26], for example, included 9 poses and 64 lighting settings per subject. In other cases (such as the “Pgu-Face” dataset [18]), objects are placed in front of the subjects to facilitate the testing of the recognition of partially occluded faces. In prior work [22,23], subjects were imaged from multiple camera perspectives and with lighting in different positions and at different levels of brightness and at different temperatures.

The dataset described herein includes multiple lighting levels; however, the subject is asked to reposition his head into multiple positions (shown in Figure 3) instead of changing the position of the camera, the lighting location or other variables. In addition, this dataset includes continuous recordings of videos of subjects’ movement, allowing assessment of recognition in the fixed positions as well as in intermediate positions.

A final aspect of datasets that should be compared is their resolution. Datasets vary significantly in this regard, ranging from smaller images such as the Georgia Tech database (with a resolution of  $640 \times 640$  pixels) [24] to multi-megapixel images (such as those presented in [22,23]). The 4 K video files presented in this dataset have a resolution of  $3840 \times 2160$  pixels.

**Author Contributions:** Conceptualization, C.G. and J.S.; data curation, C.G.; writing—original draft preparation, C.G. and J.S.; writing—review and editing, J.S.; supervision, J.S.; project administration, J.S.; funding acquisition, J.S.

**Funding:** The collection of this data was supported by the United States National Science Foundation (NSF award # 1757659).

**Acknowledgments:** Thanks is given to William Clemons and Marco Colasito who aided in the collection of this data. Facilities and some equipment used for the collection of this data were provided by the North Dakota State University Institute for Cyber Security Education and Research and the North Dakota State University Department of Computer Science.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Denimarck, P.; Bellis, D.; McAllister, C. Biometric System and Method for Identifying a Customer upon Entering a Retail Establishment. U.S. Patent Application No. US2003/0018522 A1, 23 January 2003.
2. Lu, D.; Kiewit, D.A.; Zhang, J. Market Research Method and System for Collecting Retail Store and Shopper Market Research Data. U.S. Patent US005331544A, 19 July 1994.
3. Kail, K.J.; Williams, C.B.; Kail Richard, L. Access Control System with RFID and Biometric Facial Recognition. U.S. Patent Application No. US2007/0252001 A1, 1 November 2007.
4. Introna, L.; Wood, D. Picturing algorithmic surveillance: The politics of facial recognition systems. *Surveill. Soc.* **2004**, *2*, 177–198. [\[CrossRef\]](#)
5. Wiskott, L.; Fellous, J.-M.; Krüger, N.; von der Malsburg, C. Face Recognition and Gender Determination. In Proceedings of the International Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, 26–28 June 1995; pp. 92–97.
6. Ramesha, K.; Raja, K.B.; Venugopal, K.R.; Patnaik, L.M. Feature Extraction based Face Recognition, Gender and Age Classification. *Int. J. Comput. Sci. Eng.* **2010**, *2*, 14–23.

7. Yeasin, M.; Sharma, R.; Yeasin, M.; Member, S.; Bulot, B. Recognition of facial expressions and measurement of levels of interest from video. *IEEE Trans. Multimed.* **2006**, *8*, 500–508. [CrossRef]
8. Michel, P.; El Kaliouby, R. Real time facial expression recognition in video using support vector machines. In Proceedings of the 5th International Conference on Multimodal Interfaces-ICMI '03, Vancouver, BC, Canada, 5–7 November 2003; ACM Press: New York, NY, USA, 2003; p. 258.
9. Cohen, I.; Sebe, N.; Garg, A.; Lew, M.S.; Huang, T.S. Facial expression recognition from video sequences. In Proceedings of the IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland, 26–29 August 2002; pp. 121–124.
10. Littlewort, G.; Bartlett, M.S.; Fasel, I.; Susskind, J.; Movellan, J. Dynamics of facial expression extracted automatically from video. *Image Vis. Comput.* **2006**, *24*, 615–625. [CrossRef]
11. Uddin, M.; Lee, J.; Kim, T.-S. An enhanced independent component-based human facial expression recognition from video. *IEEE Trans. Consum. Electron.* **2009**, *55*, 2216–2224. [CrossRef]
12. Valstar, M.; Pantic, M.; Patras, I. Motion history for facial action detection in video. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics—Cover (IEEE Cat. No.04CH37583), The Hague, The Netherlands, 10–13 October 2004; pp. 635–640.
13. Matta, F.; Dugelay, J.-L. Person recognition using facial video information: A state of the art. *J. Vis. Lang. Comput.* **2009**, *20*, 180–187. [CrossRef]
14. Gorodnichy, D.O. Facial Recognition in Video. In *Audio- and Video-Based Biometric Person Authentication*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 505–514.
15. The Database of Faces. Available online: <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html> (accessed on 29 January 2019).
16. Guo, Y.; Zhang, L.; Hu, Y.; He, X.; Gao, J. MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition. *arXiv*. 2016. Available online: <https://arxiv.org/abs/1607.08221> (accessed on 15 July 2019).
17. Learned-Miller, E.; Huang, G.B.; RoyChowdhury, A.; Li, H.; Hua, G. Labeled Faces in the Wild: A Survey. In *Advances in Face Detection and Facial Image Analysis*; Kawulok, M., Celebi, M.E., Eds.; Springer: Basel, Switzerland, 2016.
18. Salari, S.R.; Rostami, H. Pgu-Face: A dataset of partially covered facial images. *Data Brief* **2016**, *9*, 288–291. [CrossRef]
19. Ahonen, T.; Rahtu, E.; Ojansivu, V.; Heikkila, J. Recognition of blurred faces using Local Phase Quantization. In Proceedings of the 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008; pp. 1–4.
20. Klare, B.; Klein, B.; Taborsky, E.; Blanton, A.; Cheney, J.; Allen, K.E.; Grother, P.; Mah, A.; Jain, A.K.; Burge, M.; et al. Pushing the Frontiers of Unconstrained Face Detection and Recognition: IARPA Janus Benchmark A. In Proceedings of the Computer Vision and Pattern Recognition Conference, Boston, MA, USA, 7–12 June 2015; pp. 1931–1939.
21. Lucey, P.; Cohn, J.F.; Kanade, T.; Saragih, J.; Ambadar, Z.; Matthews, I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, San Francisco, CA, USA, 13–18 June 2010; pp. 94–101.
22. Gros, C.; Straub, J. Human face images from multiple perspectives with lighting from multiple directions with no occlusion, glasses and hat. *Data Brief* **2018**, *22*, 522–529. [CrossRef]
23. Gros, C.; Straub, J. A Dataset for Comparing Mirrored and Non-Mirrored Male Bust Images for Facial Recognition. *Data* **2019**, *4*, 26. [CrossRef]
24. Nefian, A.V. Georgia Tech Face Database. Available online: [http://www.anefian.com/research/face\\_reco.htm](http://www.anefian.com/research/face_reco.htm) (accessed on 15 July 2019).
25. Martinez, A.M. AR Face Database Webpage. Available online: <http://www2.ece.ohio-state.edu/~jaleix/ARdatabase.html> (accessed on 22 August 2019).
26. Georgiades, A.S.; Belhumeur, P.N.; Kriegman, D.J. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 643–660. [CrossRef]

