

BTAS 2016 Video Person Recognition Evaluation
Call For Participation
<http://www.pasc-eval.org/>

January 21, 2016

Patrick J. Flynn – The University of Notre Dame
P. Jonathon Phillips – NIST
Walter J. Scheirer – The University of Notre Dame

This call for participation describes the *Video Person Recognition Evaluation* being organized in conjunction with BTAS 2016. This evaluation is a follow-on to the *IJCB 2014 Handheld Video Face and Person Recognition Competition* and *FG 2015 Video Person Recognition Evaluation*. Below we summarize the task, evaluation, final report, schedule, protocol and coordination. We hope all labs and groups with active research on video face recognition will consider participating. For questions about this evaluation, please email the organizers at pascBTAS2016@pasc-eval.org.

Task: Face and Person Recognition of Moving People in Video

This evaluation will address the problem of recognizing individuals in videos. The individuals in the videos are carrying out actions such as picking up an object or blowing bubbles; generally, they are observed by the camera and the camera is not the individuals' center of attention. The evaluation emphasizes complicating factors in video taken by people using common handheld devices in everyday settings. It is assumed most approaches will emphasize face recognition, but in general all or most of the people are in full view and innovative approaches may use visual cues beyond just the face. Since there is human recognition performance for the data, we will compare human and algorithm accuracy.

Data: The Video Portion of the Point-and-shoot Face Recognition Challenge (PaSC) and The Video Database of Moving Faces and People (VDMFP)

The videos in the BTAS 2016 Video Person Recognition Evaluation are from the Point-and-Shoot Face Recognition Challenge (PaSC) and the Video Database of Moving Faces and People (VDMFP). The PaSC is summarized in the paper: "*The Challenge of Face Recognition from Digital Point-and-Shoot Cameras*" presented at

BTAS 2013. A summary of the challenge is on the PaSC website¹. The VDMFP is summarized in O'Toole et al 2005².

The PaSC videos are from the control and handheld video portions of PaSC. The control videos are from a tripod mounted video camera. The handheld video portion consists of 1401 videos of 265 people acquired at the University of Notre Dame using five different handheld video cameras. Videos are acquired at six locations: a mix of indoor and outdoor settings.

The VDMFP videos were collected in two scenarios. In the first a subject walks towards the camera. In the second, the subject to be recognized is talking with another person. The camera is looking down on conversation. We have extensive human performance data on the VDMFP videos³. VDMFP videos were included in the Multiple Biometric Grand Challenge (MBGC).

Participants are required to license both the PaSC and VDMFP data sets. There is a separate license for each data set. Both data sets are licensed from the University of Notre Dame. Download information can be obtained by going to <http://www.nd.edu/~cvrl> and following the link to "Data Sets".

- Information about the PaSC data set is found at the item entitled *The Point and Shoot Face and Person Recognition Challenge (PaSC)*
- Information about the VDMFP data set is found at the item entitled *UTD data collection*.

Each of these data sets requires execution of a license agreement by a legally authorized person at the participant institution. The VDMFP data also requires execution of a separate permission form. Follow the instructions at the top of the CVRL data sets page closely, as the agreements must be executed properly before we can grant access to the data. It can take several days to process licensing agreements.

Evaluation: Verification on Video-to-Video on Control and Handheld Video

To simplify analysis, participants' biometric matchers will be required to generate similarity scores (a larger value indicates greater similarity). If a participant's matcher generates a dissimilarity score instead of a similarity score, the scores

¹ <http://www.cs.colostate.edu/~vision/pasc>

² O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., Ayyad, J. H., & Abdi, H. (2005). A video database of moving faces and people. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(5), 812-816.

³ "Recognizing people from dynamic and stable faces and bodies: Dissecting identity with a fusion approach," A. J. O'Toole, P. J. Phillips, S. Weimer, D. A. Roark, J. Ayyad, R. Barwick, J. Dunlop, *Vision Research*, 51, pp 74-83, 2011.

should be negated or inverted in some way so that the resulting value is a similarity measure.

Participants in the evaluation will be provided with target and query sigsets (lists of biometric signatures or samples) for each of three verification experiments:

- PaSC Experiment 1 - Control: video-to-video verification on the tripod mounted camera's videos.
- PaSC Experiment 2 - Handheld: video-to-video verification on images from a mix of different handheld point-and-shoot video cameras.
- VDMFP Experiment 3 - Video-to-video verification from the VDMFP data set.

From the licensed data and the provided sigsets, participants should generate and submit three similarity matrices, one for Experiment 1, one for Experiment 2, and one for Experiment 3. The similarity matrices shall have N_t rows and N_q columns, where N_t and N_q are the sizes of the target and query sigsets. The (i,j) entry of a similarity matrix is the similarity score generated by the algorithm when supplied target sigset entry i as a gallery sample and query sigset entry j as a probe sample.

Participants will also be required to supply the companion ROC curve data for each similarity matrix.

For additional guidelines about allowable training and normalization of scores, see the section below on protocol.

Results on the experiments will be further divided into two categories:

1. Results from algorithms that automate face detection as part of the recognition process.
2. Results from algorithms that use machine generated eye coordinate information provided by the evaluation organizers. **Eye coordinates are not available for Experiment 3.**

To support this division, participants are required to state whether their results were obtained using the evaluation-supplied eye coordinates or not.

The second category is included to encourage participation from groups whose research experience may not include techniques for face or facial feature localization. Participants in the first category, doing their own detection and localization, will be invited to provide information on their process and optionally, should they choose, to share their face localization meta-data.

Report:

As has become common for evaluations of this sort, at least one paper will be written and submitted to BTAS 2016 summarizing the findings of the evaluation. The purpose of this summary paper is three fold. First, it will describe the scope and aims of the evaluation to the broader community. Second, it will provide, in one place, a record of how different approaches associated with different participants performed during the evaluation interval. Third, it will provide an opportunity for the organizers to report some analysis of these results across the various participants.

Performance across algorithms will be summarized in terms of ROC curves as well as a single performance value on those curves, namely verification rate at a fixed false accept rate of 0.01 and 0.001. For the PaSC experiments, participants will also have access to the ROC for the SDK 5.2.2 version of the algorithm developed by Pittsburgh Pattern Recognition. It achieves a verification rate of 0.49 and 0.38 at FAR = 0.01 on the control and handheld experiments.

Evaluations such as this provide the community the broadest service when they educate us about tradeoffs and underlying complications implicit in the task. The organizers will stress these tradeoffs in their summary remarks about the analyses, to avoid an inappropriate focus solely on single performance figures or the relative positions of ROC plots. The structure of the PaSC data and associated metadata facilitates the analysis of performance differences relative to distinct factors associated with handheld video face recognition. The experiment design that guided the data collection was organized around a set of factors including camera manufacturer and model, location, activity, and subject attributes. It is plausible to expect each of these experimental variables to influence performance, and to different degrees. As part of the analysis of the evaluation, performance relative to these factors will be examined. This metadata will also be made available to participants.

Beyond the summary report, it is expected that many participants will write up their own efforts and submit these for publication, hopefully to BTAS 2016.

Below is the schedule for the evaluations. Items requiring actions on the part of the participants are in boldface. Note the last day to deliver results is May 4, 2016. Also, be aware the report summarizing the evaluation is reviewed as a submission to BTAS 2016.

Evaluation Announcement.	February 1, 2016
First round similarity matrices delivered	April 15, 2016
Summary of approach description given to Notre Dame	
Final similarity matrices delivered to Notre Dame	May 4, 2015
Option to supply modified approach description	
Updated report delivered to BTAS 2016	May 10, 2016
Final notification of BTAS 2016 decision on report	June 10, 2016

We ask that all those wishing to participate please email the organizers at pascBTAS2016@pasc-eval.org as soon as is convenient. Further, groups who've not yet been through the process of preparing and delivering similarity matrices should work with Notre Dame to test that Notre Dame is able to correctly read their files. We will work with groups and try to help when necessary in order to see to it that what they generate is correctly formatted. To do this, groups need to get University Notre Dame at least one example file ahead of the April 15, 2016 deadline.

Protocol:

The evaluation will be conducted according to the PaSC protocol, which in particular requires that the similarity score $s(q,t)$ returned by an algorithm for query image/video q and target image/video t may not in any way change or be influenced by the other images in the target and query sets. This protocol therefore requires that training, as well as steps such as cohort normalization, use a disjoint set of images/videos. Here disjoint means that there are NO subjects (people) in common between imagery in the PaSC target and query sets and any training or cohort normalization sets used by an algorithm. Also, to test generalization to new locations, the protocol prohibits training on any imagery collected at the University of Notre Dame during the Spring 2011 semester and videos or images that are part of the VDMFP data set.

Training data is supplied as part of the PaSC. This training data was collected by the University of Notre Dame under circumstances broadly similar to the data in the evaluation. However, the training data comes from collection efforts carried out in semesters different from the evaluation data, and there are differences. Participants are also welcome to train on other imagery they may have available to them so long as doing so does not violate the protocol described above.

To aid participants with the details associated with running an experiment, a software package is available for participants that illustrates the complete process of running algorithms on these two experiments. The software includes a baseline video-to-video matching algorithm along with all the surrounding support code

needed to encapsulate an experiment and carry it through to the stage of writing out a similarity matrix. This software is available as part of the PaSC download⁴.

Coordination:

The organizers will encourage information sharing among participants as a means of smoothing over the inevitable myriad details that arise when working with a new dataset. Toward that end, the organizers will establish an email group and send out periodic messages. Participants are not expected to share substantive details about their own efforts during the evaluation.

The list of participants will be shared during the evaluation to facilitate communication. Any group wishing to formally withdraw from the evaluation during the evaluation may do so and mention of their participation will be avoided from that point forward. As part of the process of submitting final similarity matrices, participants will be asked to formally agree that University of Notre Dame may publicly disseminate similarity matrices through the evaluation website as well as to identify participants results by name.

⁴ http://www.cs.colostate.edu/~vision/pasc/pasc_download.php