

# IJCB 2014 Handheld Video Face and Person Recognition Competition Call For Participation

<http://www.cs.colostate.edu/~vision/pasc/ijcb2014/index.php>

November 14, 2013

*Schedule Modified December 5<sup>th</sup>, 2013*

J. Ross Beveridge – Colorado State University  
ross@cs.colostate.edu

Patrick J. Flynn – The University of Notre Dame

P. Jonathon Phillips – NIST

This call for participation describes the Handheld Video Face and Person Recognition Competition being organized in conjunction with IJCB. 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. Please direct questions to Ross Beveridge – email is shown above.

## **Task:** Face and Person Recognition of Moving People in Handheld Video

This competition will address the problem of recognizing individuals in videos captured using handheld cameras. The individuals in the videos are carrying out actions such as picking up an object or blowing bubbles; they are observed by the camera, but the camera is not the individuals' center of attention. The competition 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.

## **Data:** The Video Portion of the Point-and-shoot Face Recognition Challenge

The dataset being used for this competition is a subset of the data in the Point-and-Shoot Face Recognition Challenge (PaSC). This challenge 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 website<sup>1</sup>.

This competition will use imagery from the handheld video and still image portions of PaSC. The handheld video portion consists of 1401 videos of 265 people acquired at the University of Notre Dame using five different handheld video cameras. Frame size for the cameras range between from 640x480 up to 1280x720. Videos are acquired at six locations: a mix of indoor and outdoor settings. The still images portion of PaSC consists of images of a super-set of the people in the handheld videos. This allows for the performance of two experiments: 1) video-to-video recognition

---

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

and 2) video-to-still-frame recognition. Since the data is a subset of the PaSC dataset, participants are expected to license the data from the University of Notre Dame<sup>2</sup>. It can take several days to process licensing agreements and the instructions for executing the agreement must be followed exactly before a license can be granted.

### **Evaluation:** Verification on Video-to-Video and Video-to-Still Experiments

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

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

- Experiment 1: video-to-video verification. The target and query sigsets contain videos.
- Experiment 2: video-to-still verification. The target sigset contains still images and the query sigset contains videos.

From the licensed data and the provided sigsets, participants are required to generate and submit two similarity matrices, one for Experiment 1 and one for Experiment 2. 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 two 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 competition organizers.

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

---

<sup>2</sup> [http://www3.nd.edu/~cvrl/CVRL/Data\\_Sets.html](http://www3.nd.edu/~cvrl/CVRL/Data_Sets.html)

The second category is included in recognition of the fact that face finding and localization in the video data is itself a hard problem and our goal in organizing this competition is to encourage participation. 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 competitions such as this, at least one paper will be written and submitted to IJCB summarizing the findings of the competition. The purpose of this summary paper is three fold. First, it will describe the scope and aims of the competition to the broader community. Second, it will provide, in one place, a record of how different approaches associated with different participants performed. 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. As a point of comparison, the CSU Baseline achieves a verification rate of 0.08 at the FAR = 0.01 in both the video-to-video and video-to-still experiments. Thus, there is significant headroom and opportunity to out-perform this initial baseline. 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 0.38 at FAR = 0.01 on the video-to-video experiment and a verification rate of 0.42 at FAR=0.01 on the video-to-still experiment.

Competitions 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 competition, 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 IJCB. Depending upon the pace of the competition and the timing of the availability of results from participants, the competition organizers and the organizers of IJCB will consider organizing a special session or even a workshop organized around the competition and its participants.

**Schedule:**

Competition Announcement.	November 8, 2013
Communication of intent from participants.	February 5, 2014
First round of results submitted to organizers.	March 19, 2014
Final similarity matrices sent to the organizers.	April 10, 2014
Report on competition submitted to IJCV	May 5, 2014

**Protocol:**

The competition 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.

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 competition. However, the training data comes from collection efforts carried out in semesters different from the competition 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-still/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 download<sup>3</sup>.

---

<sup>3</sup> [http://www.cs.colostate.edu/~vision/pasc/pasc\\_download.php](http://www.cs.colostate.edu/~vision/pasc/pasc_download.php)

**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 competition.

The list of competition participants may be shared by the organizers and other participants during the competition. Any group wishing to formally withdraw from the competition during the competition 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 Colorado State University may publicly disseminate similarity matrices through the competition website as well as to identify participants results by name.