

Joint Modeling of Algorithm Behavior and Image Quality for Algorithm Performance Prediction

Apurva Bedagkar-Gala
avbedagk@mail.uh.edu
Shishir K. Shah
shah@cs.uh.edu

Quantitative Imaging Laboratory,
Department of Computer Science,
University of Houston,
Houston, TX 77204-3010, USA

Estimation of an algorithm's performance given a particular input image/video is difficult for a computer but quite easy for a human observer. Humans can assess the ability of an algorithm to generate a positive outcome for a given input based on small number of observations since they can learn high level cognitive association from input-output observations. Simulation of this process can lead to automation of algorithm performance assessment and thus eventual prediction of algorithm performance given an input image. In any computer vision system, predicting the performance of a vision algorithm can prove valuable in optimizing the overall success of the application.

In this paper, we propose a framework for predicting the performance of a vision algorithm given the input image or video so as to maximize the algorithm's ability to provide the desired output. This is achieved by modeling the performance prediction process as one that accounts for the algorithm's behavioral properties as well as the quality of the algorithm's input. A prototype system is designed for optimal prediction of vision algorithm's performance given the inputs image/video quality and its application to an optimal algorithm selection process is demonstrated. This system can be considered an intelligent system that combines algorithm's input's quality with knowledge based prediction of algorithm performance.

Performance evaluation is used to obtain knowledge about algorithm's behavioral properties [3] and intelligent system design is used to apply that knowledge. The quality of a frame is expressed as a function or combination of functions of image features that represent degradations present in the video frames. Algorithm's performance is characterized by performance metrics that capture its ability to provide a desired result.

The framework shown in figure 1 contains three main modules, *quality extractor*, *performance evaluator*, and *predictor*. The quality extractor is meant to quantify the degradations in the input image that affect the vision algorithm's performance. The performance evaluator, evaluates the algorithm's performance on a set of training input data with varying levels of image quality in order to capture its behavioral properties. In essence, the performance evaluator simulates the algorithm's perception of the input image. Finally, the predictor combines information from quality extractor and performance evaluator. It provides a mechanism to automatically acquire, store and utilize knowledge about algorithm's perception of image quality.

The framework is prototyped using 3 object tracking algorithms used in video surveillance applications. During the training phase, the input video is processed using each algorithm and performance evaluation on the results provides the performance metrics. The performance metric represents the quality of algorithm's performance associated with the input. Image quality measures are also calculated for every input frame in the video. Both image quality measure and performance measures are used to design the predictor. This allows for representation and capture of needed knowledge. When the predictor encounters a new input video, the knowledge learnt during training and the image quality are used to predict each algorithm's ability to succeed, without actual algorithm execution. The one expected to achieve maximum success is selected.

The signal activity measure proposed in [4], edge entropy and structural similarity index metric (SSIM) proposed in [5] are used to quantify image quality. Multiple Object Tracker performance evaluation is used to quantify a tracker's performance in terms of performance metrics. A systematic and objective performance evaluation of the tracker's characteristics proposed in [2] is used for this purpose. The tracking algorithms used are Uniform Motion Connected Component tracking, Mean Shift tracking, and Particle filter tracking. In order to observe and learn the effect of the degradations in input data on the algorithm's performance, we use 12 videos from the INRIA data set [1].

Table 1 shows the performance of each individual tracker and the performance of the tracker selection based on prediction on the 12 test sequences. Labels CC, MS and PF represent the Connected Component

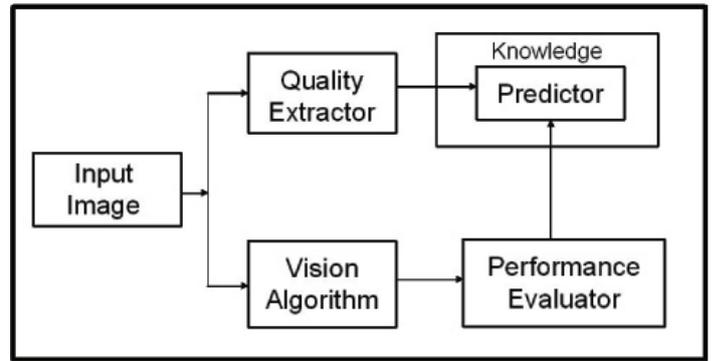


Figure 1: Performance Prediction Learning Framework.

Tracker, Mean Shift Tracker and Particle Filter Tracker, respectively. As can be seen the predictor selects the best tracker 66.6% (8/12) of the times and one of the top two trackers 91.6% (11/12) of the times.

	Rank 1	Rank 2	Rank 3	Algorithm Selection
Video 1	MS	PF	CC	MS
Video 2	CC	MS	PF	CC
Video 3	PF	MS	CC	PF
Video 4	PF	MS	CC	PF
Video 5	PF	MS	CC	PF
Video 6	PF	MS	CC	PF
Video 7	PF	MS	CC	CC
Video 8	MS	CC	PF	MS
Video 9	CC	MS	PF	MS
Video 10	PF	CC	MS	CC
Video 11	PF	CC	MS	PF
Video 12	CC	MS	PF	MS

Table 1: Predictor based Tracker Selection.

From the results, we found that the trained model for algorithm performance prediction is able to mimic the behavior of a vision algorithm relative to input quality. Such a performance prediction technique can be used to optimize the applicability of a vision algorithm to a particular application domain or serve to select an optimal algorithm from a bank of algorithms to optimize application performance. Other applications can be online algorithm switching for successful task completion or optimal parameter selection for given input in real-time. It also allows for the graceful performance degradation of a vision system.

- [1] Caviar: Context aware vision using image-based active recognition., 2004. URL <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>.
- [2] K. Bernardin and R. Stiefelhagen. Evaluating multiple object tracking performance: The clear mot metrics. *EURASIP Journal on Image and Video Processing*, 2008.
- [3] R. Haralick. Performance characterization in computer vision. In *CAIP'93*, pages 1–9, 1993.
- [4] Z. Wang, H. R. Sheikh, and A. C. Bovik. No-reference perceptual quality assessment of jpeg compressed images. In *Proceedings of IEEE 2002 International Conferencing on Image Processing*, pages 22–25, 2002.
- [5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13:600–612, 2004.