Reinforcement Learning for Active Appearance Model Dataset Selection

André Antonitsch
School of Technology
Pontifícia Universidade Católica do Rio Grande do Sul
Porto Alegre, RS, Brazil
andre.antonitsch@acad.pucrs.br

Abstract

Active Appearance Models have been used widely for facial analysis and medical image segmentation. The statistical model is based on an input dataset, and describes images based on a subset of images used for training it. Training time for the model is based on dataset size and redundancy of image features in the dataset adds little information to the model. A fast approach to select from a large dataset a good set of images to train an Active Appearance Model would facilitate the creation of datasets based on videos and speed up training times. In this paper we present a proposal of using Reinforcement Learning as a way to select a smaller dataset which should be offer a good trade off between model expressiveness and training time.

Introduction

Active Appearance Models (AAM) were created by Cootes et al. (Cootes, Edwards, and Taylor 1998; 2001), originally to fit a synthetic face onto a target face. This statistical model was shown to be useful in several tasks, such as: modeling faces, facial detection, object detection and image segmentation. While AAMs are powerful models for facial modeling and detection, their expressiveness are extremely tied to the variance contained in the dataset used to train it. This lack of generality comes from the fact that the model assumes the training dataset to be a contained in a space of what is a valid representation of the object it models (Cootes et al. 1995). Additionally, the training times for AAMs are directly connected with the dataset size used to train it.

Selecting a subset of images to use from a given dataset is a similar problem to choosing which columns are more relevant to index in a database. Using reinforcement learning for database tuning was shown by Basu et al (Basu et al. 2016).

In this project we propose using Q-Learning for dataset selection and pruning for use with AAMs. More over, we propose a linearly interpolated as shown by Tsitsiklis et al.(Tsitsiklis and Van Roy 1997). And plan on evaluating whether these techniques are capable of reducing training times for AAMs while providing minimal expressiveness with a reduced dataset.

Technical Approach

This section describes both techniques we plan on exploring, Active Appearance Models and Reinforcement Learning, as well as giving details on how we plan to evaluate the produced model.

Active Appearance Models

Interpretation-through-synthesis is an approach for analysis where the recreation of an object, or subject, through the model can lead to higher level interpretations of the features of said object. The Active Appearance Model proposed by Cootes et al. (Cootes, Edwards, and Taylor 1998; 2001) is one of such models. The AAM can be used for image analysis through the reconstruction of subjects in images, the original work presented was focused in the reconstruction and detection of faces, but its uses extend further, such as image segmentation in medical images and facial recognition.

The AAM combines a model of shape variation and texture variation in images annotated with relevant corresponding landmarks, this work focuses on the shape variation aspect of the model. For example, to build a model for faces, as we propose on this project, one might annotate the outlines of eyes, mouth, face and nose, the method assumes certain landmarks move together as the object moves (Cootes, Edwards, and Taylor 1998; 2001). The annotated points of each image are aligned in a common co-ordinate frame and each can be described by a $2n$ dimensional shape vector x, where n is the number of landmarks in the image. We then compute the deviation to the mean for each shape vector x, and build a $2n \times 2n$ co-variance matrix of the shape deviation vectors. By applying PCA to the co-variance matrix, we can approximate any example in the training set to the equation 1:

$$x = \bar{x} + P_s b_s, \tag{1}$$

where $\bar{x}$ is the mean shape vector in the dataset, P_s is the set of orthogonal modes of variation obtained and b_s is a set of shape parameters.

The model assumes the landmark points contained in the dataset define an $2n$ dimensional *Allowable Shape Domain* (Cootes et al. 1995). This space describes a region in the where each point is assumed to be a valid and possible shape. This space is assumed to be ellipsoid.

This model's expressiveness is then directly tied to how varied the dataset is, given more dissimilar points cover a larger *Allowed Shape Domain*, and therefore represent a larger possible space of valid examples. The model's computational cost also increases with the quantity of training images, therefore the use of redundant images both increases its training time, and does not add to the expressiveness of the model.

Reinforcement Learning

Reinforcement learning is an approach to how agents can learn optimal policies for action taking in a given environment (Sutton and Barto 2018; Kaelbling, Littman, and Moore 1996). Generally, reinforcement learning problems are modeled as Markov Decision Processes (MDP), where a given agent tries maximizing a their accumulated reward given a certain reward function. For this project we plan using the Q-Learning algorithm for policy learning.

Q-Learning maps an action for each possible state of the state space, this is done by iteratively applying an update function to the current mapping, whenever the agent makes an action and observes a certain result. The algorithm then maps a certain reward for taking an action a in a given environment state s as a value pair state-action, usually in a look-up table. The main advantage of this algorithm is the possibility of use in an environment where the reward function is unknown, but measurable in someway.

Ideally, Q-Learning will map every possible state-action value pair into a reward value. However, for environments which are too large or too costly to map exhaustively, alternatives have been proposed. Instead of using a look-up table, state-action value pairs can be a linear interpolation of a given feature set of a given state(Tsitsiklis and Van Roy 1997). The update step instead of updating the value associated with the value pair in the look-up table, updates the weights given to each feature in the state's feature set.

Proposed Method and Evaluation

For this project we propose a method of selecting a subset of images in a dataset which offers both acceptable expressiveness for an AAM model as well as fast training times. This speed up is expected to be particularly noticeable when using datasets extracted from videos, since those have sequential images, which are inherently redundant. To achieve this goal, we propose using a linearly interpolated Q-learning approach for dataset value representation. Given a sub-set s of the complete dataset containing images $\{i_n, i_j ... i_z\}$ we can construct a vector representing this sub-set by concatenating the landmarks of each image into a feature vector f, e.g. $\{i_{nx}, i_{jx} ... i_{zx}, i_{ny}, i_{jy} ... i_{zy}\}$.

And construct a linear interpolation function for feature vectors using a weights vector w. So a value v is associated to each sub-set s following the equation 2:

$$v = f \cdot w^T \tag{2}$$

Considering the AAM model is a linear model, it is worth investigating whether the linear interpolation function is enough or we should use a more complex interpolation function.

The evaluation for the method will be based on training time of the AAM associated with how expressive the AAM is. Expressiveness can be measured by measuring the error produced by the model over a testing dataset, alternatively, we can count how many relevant modes of variation the model produces. Testing the model over a dataset could be done with AAM matching (Cootes, Edwards, and Taylor 1998).

Project Management

We already have an AAM implementation in C++. An annotated dataset is needed. The Q-Learning portion of the method, automation of testing and dataset construction, as well as evaluation are still missing. To accomplish those tasks we propose the following schedule:

- 10/24 and 10/31 - Research into available datasets. Research into available Q-Learning implementations or frameworks. Q-Learning implementation.
- 11/07 - AAM and Q-Learning integration and automation.
- 11/14 - Model evaluation.
- 11/21 - Final assignment paper writing and presentation slides creation.

Conclusion

In this work we propose a dataset selection algorithm, specifically crafted for Active Appearance Model construction. For this, we plan on using a Reinforcement Learning approach, a Q-Learning algorithm. We will investigate whether a linear interpolation function based Q-Learning approach is adequate for this task.

If accomplished, this task will provide reduced AAM training times with minimal loss of model expressiveness. Additionally, it provides potentially huge speed ups in training time when using datasets constructed based on video data, since these contain a large portion of redundant frame data.

Acknowledgements

We thank Gabriel Licks for insights into which models to investigate.

References

Basu, D.; Lin, Q.; Chen, W.; Vo, H. T.; Yuan, Z.; Senellart, P.; and Bressan, S. 2016. Regularized cost-model oblivious database tuning with reinforcement learning. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXVIII*. Springer. 96–132.

Cootes, T. F.; Taylor, C. J.; Cooper, D. H.; and Graham, J. 1995. Active shape models-their training and application. *Computer vision and image understanding* 61(1):38–59.

Cootes, T. F.; Edwards, G. J.; and Taylor, C. J. 1998. Active appearance models. In *European conference on computer vision*, 484–498. Springer.

Cootes, T. F.; Edwards, G. J.; and Taylor, C. J. 2001. Active appearance models. *IEEE Transactions on Pattern Analysis & Machine Intelligence* (6):681–685.

Kaelbling, L. P.; Littman, M. L.; and Moore, A. W. 1996. Reinforcement learning: A survey. *Journal of artificial intelligence research* 4:237–285.

Sutton, R. S., and Barto, A. G. 2018. *Reinforcement learning: An introduction*.

Tsitsiklis, J. N., and Van Roy, B. 1997. Analysis of temporal-diffference learning with function approximation. In *Advances in neural information processing systems*, 1075–1081.