University of Regina

SOFTWARE SYSTEMS ENGINEERING

ENSE 480

Applied Artificial Intelligence

Author:

Julien POPA-LIESZ

Student ID:

200285869

January 20, 2016

Contents

1 3-D Facial Landmark Localisation

This topic explores a method for automatic localization of facial landmarks where an algorithm generates a set of candidate locations from feature detectors and performs a combinatorial search constrained by a flexible shape model. The application for this research are in the field of biometric identification and medicine. In the field of medicine such applications are facial surgery (Sharifi, 2008), lip movement assessment (Popat, 2013) or cranial facial dysmorphology. It is stated in the paper that with the increased use of 3-D scanners the possibility to overcome traditional limitations of 2-D scanners have become discovered. However according to Phillips current 3-D scanners suffer from two shortcomings: 1) not all 3-D scanners provide texture and, even when they do it cannot be assured that it is accurately registered to the geometry. 2) they may become more sensitive to view point and lighting conditions, where texture information is not invariant to these factors. (Phillips, 2005) In the presence of obstacles obstructing landmarks, for example, scarf or jacket in the way of a chin the use of shape regression with incomplete local features (SRILF) was used for the detection of such facial landmarks. SRILF handles any combination of missing points and allows for nonrigid deformations. SRILF algorithm has three components: 1) selection of candidates through local feature detection; 2) partial set matching to infer missing landmarks by regression; and 3) combinatorial search which integrates the other two components. Local feature detection is described by a set of vertices in M where M is a facial surface described by the V vertices. Then a set of 3-D annotations containing L landmarks are used to train a local descriptor model for each landmark. The goal is to compute a similarity score s(v) based on the local descriptors, that correlates well with the correct position of the targeted landmark. The second step is Partial set matching using statistical shape models. A

shape vector is constructed by concatenating the coordinates of L landmarks. The goal is to infer the coordinates of landmarks using principle component analysis over a FRGC (Facial Recognition Grand Challenge) training set. Combinational feature matching and the use of RANSAC is a matching procedure where L sets of candidate points for each landmark have a set of top scoring vertices which were determined during training period and then evaluated. The SRILF algorithm is described in the paper as follows: Start with the mesh M, for all the landmarks compute the descriptor scores, determine the landmark candidates. Next for all combinations of candidates obtain the best candidates within M. Keep iterating the subset until it achieves the desired highest score and then evaluate the geometry. In conclusion, the algorithm generates sets of candidate points from geometric cues extracted by using the APSC descriptors and then performs a combinatorial search using a constrained flexible shape model. In the scenario were a landmark was not accurately detectable the use of partial subsets of landmarks were taken to make a inference on the missing landmark. After Evaluating the facial recognition grand challenge data, errors in the ranges of approximately 3.5 mm was obtained. Such an accuracy is amazing. It is apparent that the trained data sets using artificial intelligent techniques was used but the type of technique was not stated in the paper. However in this example, applied artificial intelligence was definitely applicable in transactions systems of facial recognition.

2 Filtering Reviews by Random Individual Error

This paper tackles the problem of physician ratings websites by determining the quality of information extracted by applying artificial intelligence to detect the random individual error within the reviews. The paper goes on to describe the challenges as (1) the content and sentiment analysis of the review texts and (2) the removal of the random individual error contained in them (Michaela, Frederik, Sabine, & Valentina, 2015). To tackle these problems these researchers assigned polarities to automatically recognise opinion phrases in reviews and then check for any divergence in the rating and text. Application of this research paper aids in the current increase in fraudulent internet reviews. This is a big issue particularly in online retail, and has attracted attention in recent years from businesses and research community (Mukherjee, Venkataraman, Liu, & Glance, 2013). In some cases reviews were biased and or genuine and may appear to be deceptive which made detection of such reviews difficult in the paper. The data set used was a set if ready-made corpora physician reviews from sites such as jamede.de and docinsider.de. The paper then describes how they detect the random individual errors in the reviews by "divergent polarities" broken into 5 categories assurance, reliability, responsiveness,tangibility and time. For example some polarities can be deduced in a rate my professor situation, where a student will use words such as bad, good or incompetent too much work etc. These words can be correlated as divergent polarities. However some contradictions may occur when the review may be overall satisfactory and based on past experiences and or trust. Machine learning cannot be applied to this scenario but the approach of the application of A.I in this paper is data processing of language and error calculation from the qualitative part in the written review and the quantitative part of the actual valued result. When performing the calculations

the 5 categories are broken into subcategories for each to determine a probability of individual inconsistency in the data sets obtained from the review. Assurance is broken down into trustworthiness, kindness, consideration. Reliability can be broken down into health, education. Responsiveness is sub-categorised into waiting time. Tangibility would have a subcategory of entertainment, and time would be the time taken. In the paper they calculated the divergence of these categories by determining the arithmetic mean for each category based each random individual error per sub category. This is evaluated rate of individual error in reviews depended on the grammar annotated by the person doing the review. Training sets were generated based on this data where they compared true positives as correct and false positive erroneous reviews. They calculated the precision in ratio between true positive and false positive reviews to generate a score to compare with physician reviews. As a result the average precision of each category was 8 percent for assurance, 3 percent for reliability, 50 percent for responsiveness, 28 percent for tangibility and 14 percent for time. What was interesting was the wait time and trust evaluation results were show to be the lowest which makes sense as the data obtained from the review would be difficult to infer from individual error correlated from grammar based analysis.

3   The Regulation of Steam Pressure in a Drum Boiler by Neural Network and System Identification Technique

The application in Artificial intelligence of neural networks is applied to a complex drum boiler system in an attempt to control the output of the system in terms of steam pressure. The inputs of the system are defined as feed water flow rate with respect to heat in order to maintain a desired steam pressure inside the parameters of the boiler. In the paper the ANN method used is a training method based on Levenberg-Marquardt back propagation where the optimal model is obtained by determining the output error using a mean square when the results are filtered. With respect to ANN being applied in similar situations of a complex dynamic system where (Vasickaninova, Bakosova, Meszaros, & Klemes, 2011) used a ANN to evaluate an optimal input for a heat exchanger and compared to a PID controller. The purpose of the Levenberg-Marquardt back-propagation was to account for the non linearity of the system associated with differential evolutions. The back propagation of this type was explored by (Wu & Tsai, 2011) as means to determine the output of a speaker system. All these models have a relationship between input and their respective outputs based on some type of parameters that can be adjusted by a neural network. The model of this experiment is defined by the inputs and outputs of a dynamic system in a differential equation. The ANN algorithm is used solve this boiler non-linear problem by fitting the regression line using a sum of squares error derivation. The neural network ran for

4 Music Emotion Recognition With Standard And Melodic Audio Features

This paper explores the influence of music in the emotions of humans and how this can be studied through datasets of melodies from several songs in an attempt to research a way in searching digital music in a more advanced, flexible way. The purpose of this project is the need for a new technological innovation for the digital era of music. This is due to the increase in demand for more powerful methods of retrieving songs in the context defined by the user. There is several papers related to emotional analysis but it is apparent in this papers perspective that this idea of understanding the emotions invoked from music is a interesting approach in searching for music. There is a current standard for how music is retrieved known as "Music Information Retrieval Evaluation eXchange (MIREX) where the best algorithm used was based on system called a "mood task" were a dataset was organised into 5 types of emotional categories and obtained a search accuracy of 67.8 percent. The paper goes on to purpose an alternative approach to this model called (MER) music emotion recognition in audio by combining both the standard and melodic audio features in music. There are several models that classify emotional states in dimensions, in particular (Russell, 1980) model is the most adopted model where there are known from the basis of anger, disgust, fear, happiness, sadness and surprise. The method used for the data acquisition in this paper was to use the MIREX mood classification task to model the emotions in 5 types of clusters. Cluster 1: passionate, rousing, confident, boisterous rowdy; Cluster 2: rollicking, cheerful, fun, sweet, aim able/good natured; Cluster 3: literate, poignant, wistful, bitter sweet, autumnal, brooding; Cluster 4: humorous, silly, campy, quirky, whimsical, witty, wry; Cluster 5: aggressive, fiery, tense/anxious, intense volatile, visceral. Ultimately this data would be trained to read the wave

forms of the songs and perform further emotional analysis to better classify where the songs fit in the 5 clusters. This was done through several parameters listed as, timing, dynamics, timbre, articulation, interval, melody, harmony, tonality, rhythm, mode, loudness, musical form and vibrato. The paper goes to state that these parameters are often difficult to extract from audio signals. The use of machine learning was applied in MER under supervised learning. The goal was to supervise the machine to predict which class the test sample of the song belonged to based on previous models. This models used to train was based on ones that were extracted from a ALLMusic database and were fed into the learning machine to determine it's class. The results showed that the feature vibrato had the highest recognition accuracy in search results of 57.6 percent. The key is that the approach for emotion classification in music based on the combination of both standard and melodic audio features is feasible with machine learning. With a bigger dataset and more resources I could see this perspective on search to be an applicable way for tagging music by the emotions is can invoke.

5 Evolving A Team In A First-Person Shooter Game By Using A Genetic Algorithm

This paper explores the application of artificial intelligence in video games, in particular the first person shooter Quake III arena which was released in 1999 and its source code was open to viewing in 2005. The area of A.I applied in this field was through the concepts of genetic algorithm. Applied Genetic programming has been applied to finding winning strategies for games such as backgammon, chess and robocode (Sipper, Azaria, Hauptman, & Shichel., 2007). The use of AI has been implemented even before the popularity of video games in todayâĂŹs world. With the demand of new technologies to engage the players, game designers always try to tackle new and smarter ways of designing game agents in their video games to challenge the player in a new way. The use of a genetic algorithm was applied to a team based game in Quake III known as capture the flag. The objective in the game is to capture the flag and return it to your base. You score when the enemyâĂŹs flag is successfully return to your base. Each team consists of 5 players where they collect weapons around their environment and try to kill each other and work together to get the flag which initially is located in the centre of the map. The goal of the genetic algorithm is to evolve a computer generated team to play the game to find the optimal solution given a scenario and try defeat the opponent team. The way the game was modelled for the genetic algorithm was by the use of a finite state machine. Where transitions between states would determine a fitness score to be correlated by the genetic algorithm in order to score it self on how well it approaches a solution for a given situation. The paper defines these states as, flag at centre, flag captured by opponent, flag dropped, flag captured by team and scored (by team or opponent). For positive state transitions the fitness score would increase and for a negative transition the fit-

ness score would decrease respectively. The idea was to generate several teams and watch them evolve to a solution. The teams that had the highest fitness score would ultimately be allowed to breed offspring. The offspring would then inherit the parent process and mutate. This would then iterate until the team would have evolved and converged to a optimal solution. The proposed strategy was implemented and test results were obtained. The game was set with a match time of 30 minutes in which each team would play against a computer-controlled team for 400 generations at most. The data found that the fitness values versus generations for various states in the evolution would increase with the number of generations in all states. The study provided an interesting approach using a genetic algorithm to allow a computer team to win in a complex first person shooter environment. With such a wide search space the data was shown to converge to a solution dependant on the amount of generations. In the future with higher computing power we should expect to have more generations to increase a fitness score in a complex game environment. This will result in very intelligent game agents for users to play against.

6 Identifying Events Taking Place In Second Life Virtual Environments

The use of multi-purpose online virtual worlds is the future of human social interactions. With the increase in technology the use of virtual realities have been an outlet for most people. Second life is virtual environment developed by Linden lab. Second life is a sim based virtual reality video game that tries to emulate certain aspects of the actual world in a digital environment. The second life architecture is based on a client/server model which carries out the calculations. There are many regions in the

virtual world the user can explore. The scope of the paper is to explore the data set that second life could potentially provide in the research area of artificial intelligence. The benefits that second life can provide for AI application is the use of dynamic events that can interact with the users. The huge environment of second life allows freedom for users to work together and organise together to accomplish common goals with each other. In this respect, it is useful to have a mechanism to identify events taking place in these virtual communities in a social setting. Identifying these events taking place in the virtual training sims is useful for identifying how effective the simulation was in relation to the participates in the simulation (Cranefield & Winikoff, 2011). Cranfield highlights the importance of exploring the uses of A.I application inside second life to help enhance the realism of the virtual reality by creating intelligent agents inside the system to interact with the environment and with the users. This theory would fundamentally improve virtual reality in the future and provide users with a surreal experience. The data set extracted in second life would be the fuel to be consumed by the artificial intelligence to construct dynamic events to take place in the environment in relationship to the user and the game environment itself. The paper categories the data in second life into abstractions with levels of events occurring in the second life and how to represent all the data in one complete snap shot of the state of the entire system at a time. The events were categorised into primitive, contextual, and complex events. Primitive events were defined as instantaneous and occur at any point in time examples of these are chat, simple player interactions etc. Contextual events is a odd way of introducing a intermediate processing of low-level data abstracted from the virtual world, some examples are static information such as, location data, and position data. Complex events are defined as a set of many events. The idea of these abstractions is to convert them into all into state snap shot of the game which can be interpreted by an

AI. The end goal is use this data to implement intelligent virtual agents with better cognitive abilities. The more believable the better.

7 Summary 7

8 Summary 8

References

Chishyan, L., Wei-Hua, W., Ching-Tsorng, T., Choa-Hui, K., & Gorden, H. (2013). *Evolving a team in a first-person shooeter game by using a genetic algorithm.*

Cranefield, S., & Winikoff, M. (2011). *Verifying social expectations by model checking truncated paths. journal of logic and computation 21: 1217 to 1256. advance access.*

Kandananond, K. (2015). *The regulation of steam pressure in a drum boiler by neural network and system identification technique.*

Michaela, G., Frederik, S. B., Sabine, S., & Valentina, S. (2015). *Filtering reviews by random individual error.*

Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). *Fake review detection classification and analysis of real and pseudo reviews tech. rep., university of illinois at chicago.*

Phillips, P. (2005). *Overview of the face recognition grand challenge, in proc. ieee comput. soc. conf. comput. vis. pattern recognit. (cvpr), vol. 1. san diego, ca, usa, 2005, pp. 947 to 954.*

Popat, H. (2013). *A geometric morphometric approach to the analysis of lip shape during speech: Development of a clinical outcome measure plos one, vol. 8, no. 2, p. e57368.*

Renato, P., Bruno, R., & Rui, P. P. (2015). *Music emotion recognition with standard and melodic audio features.*

Russell, J. A. (1980). *A circumplex model of affect. journal of personality and social psychology.*

Sharifi, A. (2008). *How accurate is model planning for orthognathic surgery? int j. oral max. surg., vol. 37, no. 12, pp. 1089 to 1093.*

Sipper, M., Azaria, Y., Hauptman, A., & Shichel., Y. (2007). *Designing an evolutionary strategizing machine for game playing and beyond. ieee transactions on systems, man, and cybernetics part c: Applications and reviews 37, 583 to 593.*

Sukno, F. M. (2015, September). *3-d facial landmark localization with asymmetry patterns and shape regression from incomplete local features.*

Surangika Ranathunga, S. C. . M. P. (2012). *Identifying events taking place in second life virtual environments, applied artificial intelligence, 26:1 to 2, 137 to 181.*

Vasickaninova, A., Bakosova, M., Meszaros, A., & Klemes, J. J. (2011). *Neural network predictive control of a heat exchange. appl. therm. eng. 31, 2094 to 2100.*

Wu, J.-D., & Tsai, Y.-J. (2011). *Speaker identification system using empirical mode decomposition and an artificial neural network. expert syst. appl. 38, 6112 to 6117.*