

A VIRTUAL ANALYSIS ON VARIOUS TECHNIQUES USING ANN WITH DATA MINING

Shweta Bhatia¹, Sweety Patel², Rupal Snehkunj³

^{1,2,3}Assistant Professor, Shri Ramkrishna Institute of Computer Education & Applied Sciences, Sarvajanic Education Society, Athwagate, Surat, India
shwetabhatia97@gmail.com, sweety.jaypatel@gmail.com

Abstract

In this paper, Firstly we discussed on monitoring the quality of video in network and proposing a tool called "VQMT" (Video Quality Measurement Tool) for automatic assessment of video quality and comparing it with MOS (mean opinion score). Secondly; author had proposed a tool called "ReGIMviZ" for video data visualization and personalization system based on semantic classification also used fuzzy logic. And lastly we focus on "SOFAIT" (SIFT and Optical flow affine image Transform) technique for face registration in video to improve action unit and its various algorithms. Here, in every system the common area is ANN architecture based on supervised learning algorithm.

1. INTRODUCTION:

Considering **data mining** (for use of various data's) as major source of experiments, **artificial neural network (ANN) based on supervised learning** we focused on comparative study on various applications built using this system.

This paper had undertake comparative study in area where ANN use are: improving action units for face registration in video using dense flow, automatic quality assessment of video sequences as well as for video data visualization and personalization based on semantic organization.

As increasing in requirement of automation in technology in overall, one will require to focus on various research areas that accomplished this. Artificial neural network is one among them that provides a powerful technique for solving many problems in area such as: pattern recognition, data analysis, motion control system, quality assessment, etc. The network model used here is "feedforward" network. The artificial neuron receives input analogues to electrical impulses the "dendrites" and the output correspond to signal is sent out from neuron known as "Axon". The signals can be changed or processed by weights similar to change occur in "synapses" that is neural to neural. Neural networks have a biologically inspired modelling capability as it based on real life behaviour, but are essentially statistical modelling tools. By processing data using data mining techniques this model will provide convenient output.

2. 1st TECHNIQUE "AUTOMATIC VIDEO QUALITY ASSESSMENT BASED ON ANN".

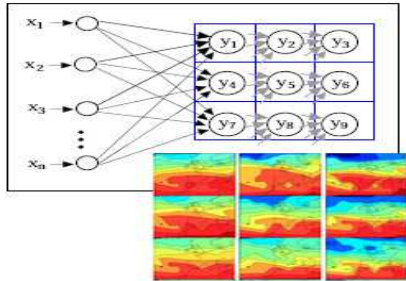
Now we discussed the first technique based on automatic quality assessment of video comparing with MOS. In reviewing this article, author had proposed the methodology for building correlation non-linear map between MOS and technical measurement to assess video quality sequences using ANN. By taking into consideration a little amount of information about original video (called reduced-reference) author wants to prove on basis of certain metric using subjective test quality of video by taking into account the human visual perception. An "VQMT" tool that had used for assessment of automatic video quality depicts 2 different application uses in multimedia n/w namely:

1. N/w monitoring and applications.
2. VQMT inside decoder with feedback of quality parameters into n/w.

Various metrics has taken into consideration for video sampling such as: PLR (Packet Loss Rate), PSNR (Peak Signal to Noise Ratio), Mean Square Error (MSE), SI (Spatial Indexes) and TI (Temporal Indexes). By taking input of quality metrics information will process and with ANN solutions rating will be available by **using VQMT tool**. On the basis on "**SUPERVISED LEARNING**" algorithm values get processed with data taken by author from database [6] and focus its experiment only considering CIF format.

Using PLR as parameter for measuring video quality, would get the resultant of packet loss rate for encountering the error occur during digital communication. The other 2 parameters

such as SI and TI are used to measure the complexity about distortion information and TI to consider how much higher value the motion in adjacent frames will take.



2.1 Experiments & Result:

The experiment results to test the ANN of 2 layers with 4 neurons (1 hidden layer) had taken with average of both EPFL [5] (i.e. database on which video experiences had carried out) and MOS, so to get closest mean human judgement of video observer’s judgement. Table 1 and 2 shows the effect of PSNR on MOS.

Table 1. Mean Opinion Score

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very Annoying

Table 2. PSNR on MOS

PSNR (db)	MOS
> 37	5 (Excellent)
31 – 37	4 (Good)
25 – 31	3 (Fair)
20 – 25	2 (Poor)
< 20	1 (Bad)

The obtained score of queried video had highly satisfied as it shows that VQMT is correlated with human observer performance (i.e. MOS), on basis of statistical parameters: spearman coefficient & Mean error between VQMT and MOS.

Points that not covered in this article are: Not mentioned the use of HDTV support with technical point, colour depth raw file support and up to how much fps to measure the video

quality can be fulfil. Author had only focus on CIF format, other format information/experiment not covered.

3. 2nd TECHNIQUE “INTERACTIVE VIDEO DATA VISUALIZATION SYSTEM BASED ON SEMANTIC ORGANIZATION”.

To overcome low-level video description, lack of scalability for long documents and no integration of user preferences in visualization of video data , author had proposed a tool called “ReGIMviZ”- which incorporates visualization and personalization concept of video data on high level extraction to group of documents while semantic various classes.

Author had undertaken various comparative studies about available visualization systems such as: HURST ZOOMSLIDER system, MEDIA MILL browsing tools for visualization, as well as SCHOEFFMANN “Instant video browsing” [12], [13], [14], [15], [16]. All this system had some disadvantages related to data scalability, indexing, segmentation as well as unpleasant aspect due to poor visual feedback.

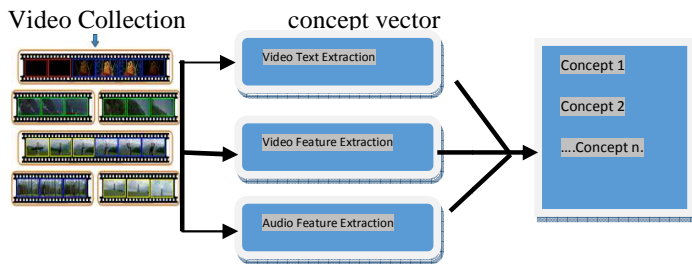
Author presented 2D Cartesian plane to display features, each axis corresponds to feature selected by user. The visualization interface is in form of neural n/w i.e. **biological neural function**.

The framework “ReGIMviZ” had used in experimenting a system that had divided into 3 parts: semantic classification, visualization and personalization. Semantic classification based on video document model and index processing. Video document model represent 2 objects: keyframe & descriptor vector.

1. The Keyframe that obtained by segmentation of video document to represent visualization space, is used to compute similarity distance between the video data.
2. Descriptor vector represents concept and context extraction.
 - a. Concept extraction used to construct vectors of video data which processes video sequences that divided into subsequence’s (shots). As shot represented by a key frame [17].
 - b. To improve similarity calculation between 2 concepts, various measurement had used such as: LCH formula, FCS deduced from NGD, etc.

An interface 2D allows visualization of data with main goal of space. This model (i.e. visualization space) is inspiration of biological neural network for represent the collection of space. The various keyframe (node) activates each other for

stimulates. Due to space limits the revelation of totality of document it should be well chosen to be posted & thus user model integration allows visualization on demand.



Personalization is based on fuzzy logic that provides greater flexibility to construct user profile. The system posts the global overview of video collection from database and after the entry of user choice, the system collects various interest centres, so to render data access process became faster by providing relevant documents and eliminating the insignificant documents that's fulfil the main goal of personalization.

3.1 Experiments & Result:

The work of video data collection here in test phase is extracted from TRECVID company database [11]. User based evaluation had been carried out by author by selecting 10 shots and the experiment based on calculated semantic similarity between documents by integrating personalization module with high precision resulted user satisfaction. While the results get more enhance by considering relation between concepts in fuzzy framework.

Points that not covered in this article are: fuzzy logic behaviour related to visualization not well represented and representation on 3D semantic classification based model support of "ReGIMviZ".

4. 3rd TECHNIQUE "IMPROVING ACTION UNIT RECOGNITION USING DENSE-FLOW FACE REGISTRATION IN VIDEO".

The goal of this system is to align faces with non-rigid muscle motion in real-world streaming video in real-time and boost facial AU (Action Unit) recognition performance. To decompose facial behaviour into possible action units (AU) one can achieved it with help of human decoders, according to "FACS" (Facial Action Coding Standard) [18]. The Challenges to fulfil in the face registration for realistic data are:

1. Rigid head motion and non-rigid muscle motion
2. Streaming data and changing resolution on face
3. The pose comprises of both in-plane rotation and out-of-plane rotation

4. The frames should comply with temporal smoothness constraint

Author had provided solution by developing an approach called "SOFAIT" that based on following:

Defined learning-based model (**SUPERVISED LEARNING**) to validate face registration, use holistic SIFT flow and optical flow based affine transform, wrap the frame to a reference with canonical pose, expression, and illumination and finally wrap the subsequent frames to its previous frame.

4.1 Technical Approach:

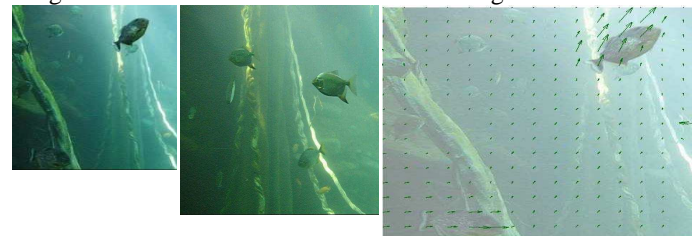
The initialization process is firstly to adopt TILT [19] - for accurately recover symmetric structure and in-plane head rotation.

SIFT flow Affine: compute SIFT flow frame to estimate for aligning faces with respect to a reference face model. This enables author system to tolerate an out-of-plane head rotation.

For temporal smoothness compute optical flow between consecutive frames, estimate the affine transform and warp the video based on the cumulated affine warping matrix.

Finally validating the current registration (i.e. error free) result using the binary classification model trained with linear SVM on HOG feature.

Original Image Database Match Wrapped and Transformed truth of original



Author had selected Avatar reference face model generated from GEMEP-FERA training dataset [21].

The TILT input face detected base on SIFT flow and Optical flow on which apply affine estimation and incorporate initialization to register the current image. Calculate HOG feature and classify the feature using trained validation model. This approach had minor the out-of-plane head rotation can be corrected by employing structural information from SIFT flow. These methods perform registration on 50 fps in real-time processing.

4.2 Experiments & Result:

Experiment had done with SOFAIT method comparing it with EAI registration approach that can prove the existing action unit system on basis of FERA and FERA2011 0challenge dataset [21], [22]. For overall AUs, author had selected temporal length parameter to generalize registration technique on per-frame basis, thus to get best FI score the parameter value consider is 0.56 second hence, 14 closest frame will be used to compute EAI representation.

Finally to get FI scores of leave-one-out cross validation, author had carried out experiment using level-1 avatar reference [20] computed from MMI, CK+ and FERA datasets.

Points that not covered in this article are The detected action units are limited and no specification of independent facial feature point detection and tracking is given.

CONCLUSIONS

In the first technique author had shown the usage of tool called "VQMT" for measuring quality of video using ANN that enables objective evolution of a given video in close correlation with Human Visual System perception. In second technique author has presented a video data visualization tool called "ReGIMviZ" that simply exploration, navigation and access of documents in large scale video corpora. The working of keyframe in this model based on ANN. And the last technique introducing a video-based real-time face registration technique that generates temporally smooth registration results on basis of dense flow-based with robustness of detecting an error, noise, etc and thus boosting the AU recognition performance. In this technique too, author had used supervised algorithm that base on ANN model for validating face registration. So our overall focus is on various techniques that are base on ANN model and experiments of every technique are based on extraction of data using data mining technique.

FURTHER ENHANCEMENT

As above all mentioned techniques can further be tested on basis of recurrent neural network (RNN).

Unlike BPTT (Backpropagation through time) that is used in above explained techniques, this algorithm is local in time but not local in space. RNN is influencing its input stream through output units connected to actuators affecting the environment.

REFERENCES:

[1]. Brice EKOBO AKOA, Emmanuel SIMEU, Fritz LEBOWSKY, "Using Artificial Neural Network for Automatic Assessment of Video Sequences".
 [2]. Jamel Slimi, Anis Ben Ammar, Adel M. Alimi, "Interactive video data Visualization system based on semantic organization"

[3]. Sofang Yang, Le An, Bir Bhanu and Ninad Thakoor, "Improving Action Units Recognition Using Dense Flow-based Face Registration in Video"
 [4]. A. Chetouni, A. Beghdadh, S. Chen and G. Mostafaoui "A novemfree reference image quality metric using neural network approach"
 [5]. F. De Simone, M. Naccan, M. Tagliasachhi, F. Dufaux, S. Turbo, T. Brahmi, "Subjective assessment of H.264/AVC video sequences transmitted over noisy channel".
 [6]. F. De Simone, M. Tagliasachhi, S. Turbo, T. Brahmi, "A H.264/AVC video database for evaluation of quality metrics". In proceeding of IEEE conf. on signal processing.
 [7]. ITU-T Recommendation P.910, "Subjective video quality assessment method for multimedia application".
 [8]. D.M. Chandler and S.S. Hemani, "Online supplement to visual signal-to-noise ratio for natural images based on near threshold and super threshold vision" 2007.
 [9]. N. Ponomarenko, V. Lukin, K. Egiazarian, Senior Member, Senior Members, IEEE J. Astola, fellow IEEE M. Charli, Senior Members, "Color image Database for Evolution of image Quality metrics", inc. Workshop on multimedia Signal Processing.
 [10]. J. Liu and D. Liang, "A Survey of FPGA-based hardware implementation of ANN", inc. Neural network brain Vol.2.
 [11]. H. Karry, A. Wali, N. Elleuch, A.B. Ammar, M. Ellouch, "Regim at treevid2008: Higher level features extraction and video search", in TRECVID 2008.
 [12]. R. Brunelli, O. Mich and C.M. Modena, "A survey on Automatic indexing of video data", Journal of visual communication and image representation, vol.10, 1999.
 [13]. M. Campanella, R. Leonardi, and P. Migliorati. "The future-viewer visual environment for semantic characterization of video sequences". In ICIP, 2005.
 [14]. W. Hurst and P. Jarvers, "Interactive, dynamic video browsing with the zoomslider interface". 2012 IEEE International conf. On multimedia and Expo. Vol 0.
 [15]. M. Worring, C.G.M Snoek, D.C. Koelma, G.P. Nguyena and O.D. Rooji, "Lexican -based browsers for searching in news video archives".
 [16]. M. Worring, C.G.M. Snoek, O.D. Rooji, G.P. Nguyena and A.W.M Smerulders, "The mediamill semantic video search engine".
 [17]. M. Del Fabro. K. Schoeffmann, and L. Bszrmnyn, "Instant video browsing: A tool for fast non-sequential hierarchical video browsing".
 [18]. Ekman, P. Friesen, W: "Facial Action Coding System: A technique for measurement of facial movement."
 [19]. Zang, Z. Liang, X. Ganesh: "TILT Transform Invariant Low-textures." In Proc. ACCV (2010).
 [20]. Yang.S. Bhanu "Facial Expression Recognition using Emotional Avatar Image". In FG workshop on FERA-challenge (2011).
 [21]. Valstar, M. Jiang, B. Mehu, M. Pantic, "The first Facial expression recognition and Analysis challenge" In proc. FG workshop on FERA-challenge (2011).

[22]. (FERA-2011: Facial Expression Recognition and Analysis challenge.

[23]. Pantic, M. Valstar, M. Rademaker, R. Maat, “ Web based database for facial expression analysis.” In IEEE conf. on multimedia and expo.