International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

Video Based Face Recognition Approach Using Genetic Approach

Om Prakash Sinha [#1], Gargishankar verma [*2]

[#] *Deptt. Of Computer Science & Engineering , CIET, Raipur*
Chhattisgarh, Bilaspur – India

[1]*omsinha9595@gmail.com* [2]gargi1981@gmail.com

***Abstract* — Extracting high level features is an important field in video indexing and retrieving. Identifying the presence of human in video is one of these high level features, which facilitate the understanding of other aspects concerning people or the interactions between people. In this paper we are giving approach for face recognition in video.**

Keywords: ***Video Processing, Computer vision, Human detection, Face recognition.***

I. Introduction

1.1 Over the last years, video-based face recognition approaches have come more and more into focus of research. They can help to overcome the difficulties arising from changing illumination, pose, expression and temporary occlusion which cannot be resolved by frames-based approaches. The most popular holistic frame-based approach to face recognition is called eigen- faces and was introduced by Turk and Pentland (1991). It describes faces as a linear combination of principal components, i. e., the eigenvectors of the covariance matrix of the training set that are associated to the largest eigenvalues. In face recognition con-text, these are generally referred to as eigenfaces. The distance of a test sample from feature space (DFFS, Moghaddam and Pentland, 1995), i. e., the energy difference be-tween the sample and its projection, can be used to determine whether it represents a face. If so, the class with the smallest distance is selected under the restriction that this distance is not too large. Otherwise, the sample is classified as unknown[3]. This property can be used to implement automatic learning of new faces. The holistic approach, how-ever, is very sensitive to occlusion and other local variations because they can affect the whole feature vector (Ekenel and Stiefelhagen, 2006a). Ekenel and Stiefelhagen (2005) choose a local appearance-based approach in order to be independent from the detection of salient facial features and less affected by local variations. This is realized using the discrete cosine transform which allows efficient dimensionality reduction and computation. In order to account for local variations in the image, the transform is applied to local blocks. The transformation results of all blocks are fused on feature and decision level, where the former is found to perform slightly superior. The local model outperforms the global version of the APPROACH AS WELL AS other holistic models like principal component (PCA), independent component (ICA) and linear discriminant(LDA) analysis. Further details are given in Sections 2.3, 3.3 and 3.4 because this approach is fundamental in this work. Zhou et al. (2003) are concerned with face recognition in video sequences rather than still images. Instead of applying the common two-step approach of tracking followed by recognition, they develop a single-step approach which is able to do both simultaneously. To achieve this, the underlying probabilistic framework incorporates a motion equation to model a person's movement, an identity equation to model the development of this person's identity and an observation equation to relate motion and identity to each other. Sequence importance sampling (SIS) is used to propagate the joint pos-terior probability distribution of identity and motion over time. Zhou et al. Present results in two categories[4]. In one, the system is trained on one single still image per person, whereas in the second, an exemplar-based approach is used to train the system with video data. Since head poses and facial expressions change continuously rather than discretely, Lee et al. (2003) represent the appearance of a person by the means of manifolds. Since a person's appearance manifold is non-linear and complex, it is divided into disjoint pose manifolds which are connected by transition probabilities. Applying PCA to exemplars, which are extracted from training videos with k-means clustering,

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
SKRGC Since 2012
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

yields an affine plane which approximates the pose manifolds. The transition probabilities are learned from the temporal dependencies between pose manifolds in training sequences. A Bayesian probabilistic framework estimates the closest manifold to a given sequence of test samples. An iteratively computed weight mask allows to handle partial occlu-sions. In order to model person-specific appearance and dynamics, Liu and Chen (2003) train individual hidden Markov models (HMM) on eigenface image sequences. During classification, the identity of a person is determined in maximum-likelihood manner. If the likelihood difference between the top two candidates is larger than a certain threshold, the sequence is used to adapt the best candidate's model accordingly. Arandjelovic and Zisserman (2005) developed a system to retrieve faces in feature-length movies based on single or multiple query images. This implies a large variety of pose and illumination changes as well as complex background and partial occlusions. SVM-based mouth and eye detectors are used in conjunction with a gradient-based face boundary detector to perform automatic face registration. To suppress coarse variations of ambient illumination, the registered face image is band-pass filtered, resulting in a so-called signature image. Classification is based on a modified Euclidean distance between the query's and film characters' signature images. The modification of the distance measure increases the robustness against partial occlusions. Using feature films as input data as well, Sivic et al. (2005) create a person retrieval system. Single faces are represented with scale invariant (SIFT) descriptors of facial features. Face detections within the same shot are automatically grouped into facetracks, which are represented by a histogram over the corresponding faces.

1.2 Face Detection

Face detection has attracted a lot of research effort in the past, as it is the key to more sophisticated systems that allow face tracking or recognition. A large variety of different methods has been developed which can be divided into two categories: feature-based and image-based (Hjelmas and Lee, 2001). Image-based approaches directly evaluate pixel values in some way, e. g., by feeding them into a neural network. Rowley et al. (1998) used a set of rationally connected neural networks to classify input samples as face or non-face. In order to compensate for translation, a window is shifted pixelwise over the image and each location is evaluated. To account for scale variations, this approach is applied to a pyramid of stepwise downsized versions of the input image. Both the sliding window and the scaling lead to a high computational effort. As this is generally a concern for image-based approaches, it applies as well to the work of Menser and Müller (1999), which is based on the eigenfaces approach introduced by Turk and Pentland (1991) for face detection and recognition (see below, Section 1.2.2). Every subwindow is projected onto the face space by means of principal component analysis (PCA). Instead of processing the original image, Menser and Müller use a skin probability image to increase robustness towards complex backgrounds and illumination. In addition to the residual reconstruction error, the distance to the mean face in the face space is used as a measure for "faceness" which increases the robustness in uniform backgrounds. Feature-based methods exploit certain characteristics of a face. These can be lowlevel features like edges, skin color or motion, the position of eyes, nose and mouth or the geometric relations between them, for example. One representative of this class is the component-based approach by Heisele et al. (2001b), which is applied to synthetic face images derived from textured 3-dimensional head models. The system uses 15 support vector machines (SVM), one for each of the 14 components, and one to classify their geometric configuration. Each component is selected by growing a rectangular region around a seed point as long as the upper bound on the expected probability of error of the corresponding SVM decreases. This approach yields a discriminating set of components and is not restricted to faces. Heisele et al. report an increased performance over a holistic approach, because the single components are less affected by in- and out- of-plane rotations. Papageorgiou et al. (1998) introduced a general framework for object detection based on an over complete dictionary of three types of Haar wavelets. These are used to com-pute differences between image intensities in neighboring regions of the image. A statistical analysis of the normalized resulting values allowed them to reduce the number of features that are necessary for a successful detection by nearly two orders of magnitude. Classification of pre-aligned

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

faces is done with SVMs. During training, false detections in non-object samples are iteratively used to refine the model until a satisfying decision surface is found. Selection of these problematic samples for training overcomes the problem that the non-object class is extremely large and heterogeneous compared to the object class. A similar experiment was conducted for pedestrian detec-tion in which an extension incorporating motion information was able to improve the results. The idea of using Haar basis functions to compute intensity differences was taken up by Viola and Jones (2001) to build a real-time capable face detection system. They extended the set of Haar-like features and introduced a data structure called integral im-age, which allows efficient computation. Thus, faces of different sizes can be detected by scaling the features instead of the image, as done by Papageorgiou et al., which results in a speed-up. Viola and Jones used a modified AdaBoost algorithm to select relevant features. These are arranged in a classifier cascade to further speed up process-ing. In analogy to the training method used by Papageorgiou et al., Viola and Jones use false detections of one classifier to train its successor in the cascade. As this approach is widely employed throughout this work, a more detailed explanation can be found in Section 2.1. Later, Jones and Viola (2003) added a decision tree for pose estimation so that an appropriate cascade could be selected to allow for multi-view face detection. Since a detailed survey of the wide variety of face detection techniques is beyond the scope of this work, the interested reader is referred to the work of Hjelmas and Lee (2001) and Yang et al. (2002).

II. RELATED WORK

A: *Feature Identification*

Most of the studies in this field use face detection algorithm as the key idea. Jin [15] proposed a method to identify video shots with people based on face detection. The category of the shot was considered to be "people", only if there is at least one image with more than one face within that shot. One of the three features chosen by Huang et al. [16] to be evaluated in the TREC video Evaluation (2003) was "People" feature, Huang et al., state that for a segment of video to have people feature it should contain at least three human faces. Huang et al. used a skin tone filter to detect skin regions, followed by the omni-face detection algorithm which was proposed by Wei and Sethi [23].

B: *Human Detection*

From the literature reviews done, it can be concluded that most common way in human detection is via detecting human face. Human face is the most unique part in human body, and if it is accurately detected it leads to robust human existence detection.

C: *Face Detection methods*

Several studies were done in face detection field since 1970, and lots of surveys addressed the algorithms used in this field under different categories [5], [13], [14] but in general two main classes can be used to classify these algorithms namely, feature based (e.g. Bottom-Up) and image based (e.g. Appearance-Based and Template matching) approaches. Features based approaches extract facial features from an image and manipulate its parameters such as angles, size, and distances. Image base approaches rely on training and learning set of examples of objects of interest. However, dealing with video introduces other approaches for face detection such as motion based approach. A brief description of the most common approaches and examples of algorithms used in each of them is given in the rest of this section.

C1: *Knowledge-based (Top-Down) approach*

In this method the relationship between facial features is captured to represent the contents of a face and encode it as a set of rules. Coarse-to fine scale is used in lots of algorithms classified under this category, in which the coarsest scale is searched first and then proceeds with the others until the finest scale is reached.

C2: *Feature invariant (Bottom-Up) approach*

In this approach, the face's structural features which do not change under different conditions such as varying viewpoints, pose angles and/or lightning conditions. Common algorithms used under this category are: Color-based approach, or so called skin-model based approach. This approach makes use of the fact that the skin colour can be used as indication to the existence of human using the fact that different skins from different races are clustered in a single region. Cezhnevets et al., [21], presented 4 pixel-based skin modeling techniques named as Explicitly defined skin region, Non-parametric skin distribution modeling, Parametric skin distribution modeling, Dynamic skin distribution modeling.

C3: *Facial features based approach*

This method, in which global (e.g. skin, size, and shape) and/or detailed (e.g. eyes, hose, and lips) features are used, has become popular recently. Mostly, the global features first

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

are used to detect the candidate area and then tested using the detailed features. *Texture* The human face differs from other objects in texture. This method, examines the likelihood of sub image to belong to human face texture, using Space Gray Level dependency (SGLD) matrix.

C4: *Template matching methods*

These methods are based on measuring the degree of similarity between the candidate sub image and the predefined stored face pattern. The predefined image might be for the whole face pattern or the individual face features such as eyes, nose and lips. Common algorithms used under this category are: Predefined face templates, in which several tem plates for the whole, individual or both (whole and individual) parts of a face are stored. Deformable Templates in which an elastic facial feature model as a reference model where the deformable template mode of the object of interest, is fitted in.

C5**:** *Appearance-Based Method*

Unlike template matching methods, where the templates are predefined by experts, Appearance-Based method learns the templates from set of images, using statistical analysis and machine learning. Examples of algorithms used by these approaches are:

Eigenfaces, or so called eigenvectors, in which different algorithms are used to approximate the eigenvectors of the auto correlation matrix of a candidate image. [27]

Distributed-Based, where the distribution pattern of an object is learned using the positive and negative image sets of that object.

Neural Networks, where networks of neurons (simple Elements) called nodes are used to perform function in parallel. The idea of neural networks comes from the central nervous system. However, these networks are trained to detect the presence of face by giving it face and no face samples.

Support Vector Machines, these are learning machines that make binary classifications. The idea here is to maximize the margin between positive and negative sets of vectors and obtain an optimal boundary which separates the two sets of vectors. They were first suggested by Vapnik in 1960 [4].

Hidden Markov Model is a statistical model used to model the statistical properties of a signal. The Markov process is used to model the processed system and the Markov parameters are taken from the observed parameters.

D: *Movement Detection*

Unlike still images, video sequences hold more details about the history of moving objects (foreground), which help in isolating the foreground from the background. Generally, the moving areas are detected by finding the changes that happen among the sequences of images [1], [2].Most of the research done in movement detection applied pre-processing steps before applying the change detection algorithms, [2]. Such pre-processing steps involve geometric and intensity adjustments. The problem of variation in light intensity is solved by intensity

adjustment in which illumination effect is reduced to some degrees based on the method used. Elgammal et al. [1], state that transforming the RGB values, into chromatic colour space makes the module insensitive to the small changes in the illumination. There are several ways for detecting a change in a video sequence [2]. Recent studies agree that Image differencing method is more effective than others in change detection [3].

III PROPOSED METHODOLOGY

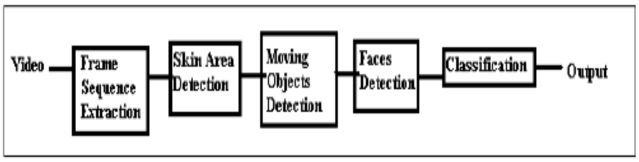

Figure 2: Steps involved in human detection

Our proposed algorithm comprises of following steps.

1. Converting a video sequence in to individual images.
2. Accessing the sequential images and detecting the important features.
3. Allocating those regions (if any) giving indications of human presence such indication is having a human skin like colour.
4. Applying movement detection test for all of the allocated regions.
5. Applying face detector to those detected moving objects to detect if it is a face or not.
6. Select a best face and add this face into Database for matching with train Database.

Stage 1: Frame Sequence Extraction

In this stage video (MPEG formatted videos) is converted into individual JPG frames

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

.Stage 2: Skin Areas Detection

In this stage colour information of the digital image is utilized to find those areas close to human skin colour. This stage helps in reducing the search space and therefore speeds up the simulation by consuming the processing time efficiently. However, skin test is not enough to detect human faces as it will also detect other parts of the body as well as other non face skin colored objects. Thus, other tests to filter out those unwanted areas should be applied. Further stages in the proposed project are designed to gradually eliminate the false detected areas found at this stage. The first test to remove the unwanted skin like areas was chosen to be movement detection.

Stage 3: Moving Objects Detection

To minimize the errors in face detection we can utilize the human nature that human will have at least small amount of movements such as eyes blinking and/or mouth and face boundary movements. We can get this information easily because we are dealing with video sequence by which the whole sequence of the object's movements can be obtained. Taking that point in to account we can reduce the error that occurs due to false detection of a human face and minimize the time of simulation. This step was designed to be implemented only across those skin regions found in the previous step. Giving those moving pixels different colour than surrounding region (human face skin colour in case of face was detect), these pixels reshape the human face facial features which in turn helps in later stages. However it is important to take in to account that a change may occurs due to several sources such as moving objects, presence or absence of objects camera movement and zooming, brightness changes This means that some changes are significant and others are not and this is varying with the application requirements. For example, the change detected in background is not significant in video surveillance whereas it has a great importance in remote sensing. Although, it is difficult to take decision whether a detected change is significant or not, it is an important step to remove unwanted changes and focus the processing only on those changes of interest, which reduced the processing time and false detected areas. Hence, movement detection was chosen as a vital stage in the proposed design.

Stage 4: Face Detection

To insure that the moving part is a face, additional tests are required. In this stage, the moving objects which were detected in the previous stage are examined to identify if any of them is a face by examining the pass of the following four tests:

Geometric Test

The candidate regions are tested here against some human face geometric features which are governed by the relation between the width and height of the human face. This test is important to eliminate some of those regions which contain non face objects whose colours are similar to the human face skin colour and experience some acceptable movement across the frames sequence.

Temporal Test

In this test the advantage of having the temporal information from the frame video sequence is being utilized to help in constructing additional verification step before applying the further tests which are more computationally expensive. The principle used in this step based on the fact that no face can occur or disappears suddenly in a certain sequence hence, comparison with the previous and next frames (if exist) gives indication whether or not the candidate region is a face. This additional step helps in maximizing the elimination of those detected skin areas that do not include a human face. Therefore, a reduction in the computational efforts as well as simulation time will be achieved.

Facial Feature Test

This stage will examine the existence of the facial features (mainly the moving areas (none skin) such as eyes, lips and face boundary) in the candidate skin areas filtered in the previous test. It is useful to use the fact that the face skin region must have at three separate spots as a test to be applied in order to verify the existence of a human face in the candidate areas. Only those candidates passing this test will be proposed for the next verification tests.

Template Matching Test

Those candidate regions that have passed the previous two tests will be compared (correlated) with a template model of a human face. Only those candidates that achieve a correlation value beyond a pre defined threshold and a distance (from the face space) value less than a pre defined threshold will be considered as a human face.

Stage 5: Classification

Only those allocated faces that have passed all of the verification stages successfully undergo the classification stage where the category of the detected candidate region is classified as either a face or not face in this stage based on the results of the last two tests namely correlation and distance from face space tests. The candidate region to be accepted as a face has to have a correlation value above the specified correlation threshold value and its distance from the face space should be below the specified distance threshold value.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

IV CONCLUSION

An algorithm has been proposed to detect the presence of human in video sequence. The main two techniques used in building the proposed algorithm are face and motion detection techniques. A series of stages were implemented in a certain order to promise maximizing the detection of existing faces and eliminating theother objects (noise). The proposed algorithm detects faces of different sizes under different lightning conditions.

V FUTURE RECOMMENDATIONS

The proposed work handles the first step in all of those applications concerning human recognition or human activities such as identifying the shots that include different activities such as meeting, running and hand shaking. Further more, its contribution in more advance applications where integration between image processes and audio processing is required to understand the videos and find out different high level features such as identifying the speaker in a certain sense, widens its importance and its applied field.

REFERENCES

[1]. Elgammal.A., Duraiswami.R., Harwood.D., And Davis.L., (2002) "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance" Proceedings of the IEEE. vol 90, no. 7 pp.1151-1163

[2]. Richard J. Radke, Andra.S., Al- Kofahi.O., and Roysam.B**, (2005)"** Image Change Detection Algorithms: A Systematic Survey". IEEE Transactions on Image Processing, VOL. 14, NO. 3.

[3].http://www.cast.uark.edu/local/brandon thesis/Chapter_IV_change.htm last accessed on [3th of September 2008]

[4]. Burbidge.R., Buxton.B., "An Introduction to Support Vector Machines for Data Mining", available from: http://www.martinsewell.com/dataminin g/BuBu.pdf

[5]. Peer .P., Solina.F., (1999). "An Automatic Human Face Detection Method", Project J2-8829, available from [http://lrv.fri.unilj. si/~peterp/publications/cvww99.pdf], last accessed on [1st of May 2007]

[6]. Kriegman.J., Yang.M., Ahuja.N., 2002 "Detecting Faces in Images: A Survey" IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 24, NO. 1 pp. 34-58

[7]. Gonzalez, Rafael C. & , Woods.R.E. (1992), Digital Image Processing /Rafael C. Gonzalez, Richard E. Woods . - Reading, Mass.; Wokingham : Addison-Wesley, (1992)

[8]. Kovac. J., Peer. P. and Solina. F., (2004) "Human Skin Colour Clustering for Face Detection" available from: [http://www.princeton.edu/ ~aabdalla/ele 79/Human%20Skin%20Colour%20Cl ustering%20for%20Face%20Detection. pdf] last accessed on [1st of September 2008]

[9]. Peer. P., Solina. F., Batagelj. B., Juvan.S., & Kovaˇc. J., (2003) "Colour-Based Face Detection in The "15 Seconds Of Fame"

[10]. Clarke, R. J. (Roger John), Digital compression of still images and video / R. J. Clarke (1995). London: Academic Press, 1995.

[11]. http://en.wikipedia.org/wiki/Vi deo_codec, last accessed on [5th of October 2008]

[12]. http://en.wikipedia.org/wiki/Vi deo_compression, last accessed on [5th of October 2007]

[13]. Yow. K. C., Cipolla. R., (1997) "Feature-based Human Face detection", no. 15, pp. 713-735.

[14]. Low, B. K & Hjelm°as1, E. (2001), "Face Detection: A Survey"

[15]. Jin, R. & Hauptmann, A. G., **"**Learning to Identify Video Shots With People Based on Face Detection", July 6-9. 2003

[16]. Huang, X., Wei, G. & Petrushin, V. A., "Shot Boundary Detection and High level Features Extraction for the TREC Video Evaluation 2003"

[17]. http://java.sun.com/products/javamedia/jmf/2.1.1 html, last accessed on [9th of September 2008]

[18]. Over, P., Ianeva, T., Kraaijz, W., & Smeaton, F. A."TRECVID 2005 – An Overview", March 27, 2006.

[19]. Kuchi, .P, Gabbur, P., Bhat, P. S., & David, S. S. "Human Face Detection and Tracking using Skin Colour Modeling and Connected Component Operators", (2002)

[20]. http://java.sun.com/products/ja va-media/jmf/, last accessed on [6th of October 2008]

[21]. Vezhnevets, V., Sazonov, V., & Andreeva, A., "A Survey on Pixel- Based Skin Colour Detection Techniques"

[22]. Gomez, G., Morales, E.F. "Automatic Feature Construction and a Simple Rule Induction Algorithm for Skin Detection"

[23]. Wei, G. & Sethi, I. K., (2000) "Omni-Face Detection for Video/Image Content Description"

[24]. ttp://www.ucl.ac.uk/MicroCore/HTML_resource/PCA_1.ht m last accessed on [29th of September 2008]

[25]. http://en.wikipedia.org rhunen-Lo%C3%A8ve_transform last accessed on [29th of October 2008]

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
Since 2012
http://www.ijcsjournal.com Volume 10, Issue 1, No 1, 2022 ISSN: 2348-6600
Reference ID: IJCS-387 PAGE NO: 2645-2651

[26]. http://en.wikipedia.org/wiki/Ei eigenface last accessed on [20th of October 2008]
[27]. M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cognitive Neuroscience,3(1), 1991a.URL http://www.cs.ucsb.edu/~mturk/Papers/j cn.pdf.
[28]. Ashish Tiwari. Face Recognition Eigen-faces with 99 PCA
coefficients.ppt
[29]. Jon Krueger, Doug Kochelek,Marshall Robinson & Matthew Escarra.
[30]. Thresholds for Eigenface Recognition. Version 1.2: Dec 17, 2004
[31]. Hsu, R., Abdel-Mottaleb, M. & jain, A. (2002) "Face Detection in Colour Images"
[32]. http://goldennumber.net/ last accessed on [9th of October 2008]
[33].
http://www.cl.cam.ac.uk/dtg/attarchive/facesataglance.html
last accessed on [9th of October 2008]
[34]. http://www.fourcc.org/fccyvrg b.php last accessed on [19th of November 2008]
[35]. http://agingmind.cns.uiuc.edu/f acedb/ last accessed on [19th of November 2008]
[36]. "Receiver operating characteristic". Available from http://en.wikipedia.org/wiki/Receiver operating characteristic, last accessed on [19th of November 2008]