

Real-Time Gender Classification by Face

Eman Fares Al Mashagba
Computer Sciences Department
Zarqa University
Zarqa, Jordan

Abstract—The identification of human beings based on their biometric body parts, such as face, fingerprint, gait, iris, and voice, plays an important role in electronic applications and has become a popular area of research in image processing. It is also one of the most successful applications of computer–human interaction and understanding. Out of all the abovementioned body parts, the face is one of most popular traits because of its unique features. In fact, individuals can process a face in a variety of ways to classify it by its identity, along with a number of other characteristics, such as gender, ethnicity, and age. Specifically, recognizing human gender is important because people respond differently according to gender. In this paper, we present a robust method that uses global geometry-based features to classify gender and identify age and human beings from video sequences. The features are extracted based on face detection using skin color segmentation and the computed geometric features of the face ellipse region. These geometric features are then used to form the face vector trajectories, which are inputted to a time delay neural network and are trained using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) function. Results show that using the suggested method with our own dataset under an unconstrained condition achieves a 100% classification rate in the training set for all application, as well as 91.2% for gender classification, 88% for age identification, and 83% for human identification in the testing set. In addition, the proposed method establishes the real-time system to be used in three applications with a simple computation for feature extraction.

Keywords—*Biometrics; Face Detection; Geometry-based; Gender Classification; Quasi-Newton Algorithms*

I. INTRODUCTION

Automatic analysis of video data is a very challenging problem. To find a particular object in a video stream and automatically decide if it belongs to a particular class, one should utilize a number of different machine learning techniques and algorithms, as well as solve object detection, tracking, and recognition tasks [12].

Automatic human recognition tasks based on pattern recognition and artificial intelligence (AI) use different biometric body parts, such as face, fingerprint, gait, iris, and voice. Of all these body parts, the face is one of the most popular traits because of its unique features and impression [9]. Identity verification using face recognition can be performed without the cooperation and knowledge of the person being identified. However, recognizing a person becomes difficult because of the variations in pose, illumination, occlusion, expression, and so forth. [9]. Human identification using the face is critical because it considers multiple parts of the facial structure, and the face of a person changes with the passage of time. In fact, individuals can process a face in a variety of ways

to categorize it by its identity, along with a number of other demographic characteristics, such as gender, ethnicity, and age. In particular, recognizing human gender is important because people respond differently depending to on gender. In sum, a successful gender classification approach can boost the performance of many other applications, including person recognition and smart human–computer interfaces [20]. Face detection is critical to the final result in several applications, such as face processing (i.e., face, expression, gender classification, and gesture recognition), computer–human interaction, human crowd surveillance, biometrics, video surveillance, AI, and content-based image retrieval. It can be viewed as a preprocessing step for obtaining the object region [7] [12] [20] [21].

Recent research carried out face detection by using a color-based algorithm [5] [7] [10] [15] [18] [22] [24], where the segmentation of skin-colored regions becomes robust only when a proper color model is chosen. Several color models exist RGB, YCbCr, and HSV color models, and each has a specific work field and strength [7]. The RGB color space consists of the three additive primaries: red, green, and blue. The RGB model simplifies the design of computer graphics systems but is not ideal for all applications. The YCbCr color space was defined in response to increasing demands for digital algorithms in handling video information and has since become a widely used model in digital videos. Given that hue, saturation, and intensity value are properties used to describe color, a corresponding color model, HSV, logically exists. When using the HSV color space, knowing what percentage of blue or green is required to produce a color is unnecessary; the hue is adjusted to get the desired color [7].

The rest of this paper is organized: Previous Studies in section II, proposed methodology in section III, Experimental results in section IV, finally conclusion and future work in section IV.

II. PREVIOUS STUDIES

The present study proposes a short review of the latest methods used in gender classification. In facial feature extraction, various methods apply to extract the features from image or video sequences: geometry-based, template-based, color-based segmentation, and appearance-based methods [6] [21] [22].

The geometry-based method extracts features using geometric information, such as the relative positions and sizes of the facial components. This method, however, requires the classifier to use a large number of features. The techniques proposed in [6] extract geometric features, such as left eye width, righteye width, nose width, left eye center to mouth left

corner, right eye center to mouth right corner, left eye center to mouth right corner, and right eye center to mouth left corner, mouth left corner to middle of chin, and mouth right corner to middle of chin. [17] proposed using geometric features, such as the distance between eyebrows to an eye, eyebrow to nose top, nose top to mouth, eye to mouth, left eye to right eye, nose width, and mouth width, all of which are extracted using the Viola–Jones algorithm. [1] [8] [15] applied a combination of global and grid features. Global features include inter-ocular distance, the distance between the lips to the nose tip, the distance between the nose tip to the line joining the two eyes, the distance between lips to the line joining the two eyes, eccentricity of the face, ratio of dimension, and width of the lips. Grid features include skin color, moustache region, lip region, eye tail, forehead, eyelid, and nose wing of the face image.

The template-based approach matches facial components to previously designed templates using an appropriate energy functional. The best match of a template in the facial image will yield the minimum energy; this approach can be too complex due to the extensive computation involved and is only effective when query and model images have the same scale, orientation, and illumination properties. Color-segmentation makes use of skin color to isolate the face and any non-skin color region. In this approach, the quality of the image coupled with illumination and hue plays an important role in image recognition and rate of classification. For a complicated background, [18] based their face detection method on the skin color feature.

The features in the appearance-based method differs from the simple facial features, such as eyes and mouth, used in other approaches. Any extracted characteristic of the image refers to a feature. In methods such as principal component analysis (PCA), which was proposed by [5] [14] [19] [21] [24] [27], a face image is represented as a two-dimensional N by N array of intensity values or a vector of dimension N^2 . Then PCA finds an M -dimensional subspace whose basis vectors correspond to the maximum variance direction in the original image space. New basis vectors define a subspace of face images called face space. All images of known faces are projected onto the face space to find sets of weights that describe the contribute of each vector. The face can be identified through a comparison of a set of weights for the unknown face and the sets of weights of known faces. [12] applied an adaptive feature generation algorithm trained by means of the optimization procedure according to the LDA principle. Gabor wavelets were also used to extract the feature vector. These approaches are commonly used for facial recognition rather than person identification. [3] [24] applied Gabor filters at five scales and eight orientations for detecting skin regions. In [13], a Fourier–Gabor filter was applied to extract features from the face images. A different size, orientation, and scale values were adopted.

Regardless of the above mentioned methods, a variety of classification techniques can be used for recognition, such as decision trees, neural networks, nearest neighbors classifier (NNC), support vector machines (SVMs), Bayesian networks, fuzzy logic, and genetic algorithms, among many others. [11] trained and tested three classifiers SVM, back propagation NNs

(BPNNs), and KNN. These classifiers are then optimized through the GA; using this approach; they got promising results in terms of the classification error rate and the minimization of computation time. [6] proposed a variant of the decision tree algorithm for gender classification of frontal images owing to its distinctive features. Their technique showed robustness and relative scale invariance for gender classification. [22] got high accuracy by using SVMs for gender classification. [3] [21] [26] produced very promising recognition rates for three applications face recognition, facial expression recognition, and gender classification, and reasonable results in all databases with the same set of features and (NNC) classifiers. The system also had real-time capability and was automatic. [23] used an unsupervised learning technique to classify DCT-based feature vectors into groups and identify if the subject in the input image was “present” or “not present” in the image database. After training for approximately 850 epochs, the system achieved a recognition rate of 81.36% for 10 consecutive trials. The main advantage of this technique is its high-speed processing capability and low computational requirements in terms of speed and memory utilization. In [14], PCA matrixes with different numbers of components (40, 50, and 60) were passed to the NN, which conducted the framework with a different layer number, different learning algorithms, and a different number of neurons in each layer, and achieved a high correct classification rate of approximately 83.5%. [8] used posteriori class probability and Artificial Neural Network to classify gender and age, respectively. The results achieved were 100% for face recognition, 98% for gender classification, and 94% for age classification. In [18], the recognition step was applied by SVM. A good performance of the gender classification test was also achieved on a relatively large-scale and low-resolution video database. [17] applied the Artificial Neural Network for face, facial expression, and gender classification, and produced reasonable results in all databases. [12] proposed using SVM and presented the experimental results gained on a large image dataset. More than 90% accuracy of the viewer's gender recognition was achieved.

The myth of a successful real-time gender classification depends greatly on the right choice of features and classification method. In this paper, an algorithm is introduced to enhance real-time gender classification. The face region is detected using skin color segmentation with HSV color space, and the features are extracted using global geometric properties. A Time Delay Neural Network (TDNN) algorithm is used for classification tasks. The suggested method is implemented by using our own datasets.

III. PROPOSED METHODOLOGY

The steps of the proposed method can be summarized as follows:

- 1) *Extract video frames.*
- 2) *Delete abnormal images from the start and the end manually.*
- 3) *Detect the face region from each frame using skin color segmentation.*

- 4) Extract features from each image using the geometry-based approach.
- 5) Classify gender using dynamic neural networks.

A. Face detection

With our method, a dependable result can be achieved for video face segmentation based on skin color. The quality of face detection is critical to the final result of the whole system as an imprecise determination of face position can lead to wrong decisions at the recognition step [12].

Each video sample is taken from our own database. The segmentation of skin-colored regions becomes robust only by choosing proper color model. Numerous applications use the HSV color model. Machine vision uses HSV color space to identify the color of different objects. Image processing applications, such as histogram operations, intensity transformations, and convolutions operate only on an intensity image. These operations are performed with much ease on an image in the HSV color space [2] [7] [12]. We applied skin color segmentation with the HSV color model to separate all skin color pixels in connecting components that use an adaptive Gaussian mixture model (GMM) proposed by [10]. The adaptive GMM can adapt the model parameters to cope with changing imaging conditions, such as lighting and noise. The segmentation method detected the ellipse that represents the skin region, both original and skin image, as shown respectively in Figure. 1(a) and (b).



Fig. 1. (a) Initial Image (b) Extracted Region

B. Feature Extraction

Geometric-based features are extracted using ellipse mathematical definition and properties [28]. Global features are calculated for a video sequence frame from the ellipse for the face region, horizontal center (X_0), and vertical center (Y_0) of the detected face, where a and b are the semi-major and semi-minor axes (half of the major and minor axes of the ellipse) respectively. The area enclosed by an ellipse area is shown as Equation (4). The two foci (the term “focal points” are also used) of an ellipse are two special points F_1 and F_2 on the ellipse's major axis and are equidistant from the center point. The sum of the distances from any point P on the ellipse to those two foci is constant and equal to the major axis ($PF_1 + PF_2 = 2a$), denoted as eccentricity as in Equation (3). The eccentricity of an ellipse, usually denoted by ϵ or e , is the ratio of the distance between the two foci to the length of the major axis or $e = 2f/2a = f/a$. The ratio of dimension and four-quadrant inverse tangent (arctangent) of the real parts of the semi-major axis (a) and semi-minor axis (b) are calculated using Equation (2). Figure 2 shows the geometry of the face region. The distance from the center C to either focus is $f = ae$,

which can be expressed in terms of the major and minor radii as shown in Equation (1):

$$f = \sqrt{a^2 - b^2} \tag{1}$$

$$P = \text{atan2}(a, b) \tag{2}$$

The eccentricity of the ellipse (commonly denoted as either e or ϵ) is

$$e = \epsilon = \sqrt{\frac{a^2 - b^2}{a^2}} = f/a \tag{3}$$

The area enclosed by an ellipse is

$$\text{area} = \pi ab \tag{4}$$

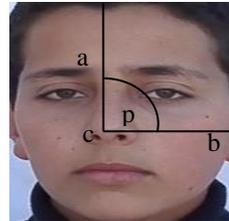


Fig. 2. Geometry of the Face Region

The features used to track the difference between each face region for given adjacent frames in the sequences and the resulting spatio-temporal vectors input in the classification step.

C. Classification Method

Finally, a TDNN is used for gender classification. The TDNN is a well-known classifier that is very successful in classifying spatio-temporal patterns as described in details in my previous research [16] and is thus used in this research. The architectural characteristics of the TDNN have an additional feature, known as a tapped delay line in its input neuron [12]. In the present research, the delays are: in the input layer, $d_1 = 0:3$; in two hidden layers, $d_2 = 0:3$ and $d_3 = 0:5$; and in the output layer, $d_4 = 0:5$. Training and testing require a sufficient number of video databases. There are many commonly used databases for the tasks of human face recognition such as Multi-PIE database, the FERET database, and SCface, but they donot contain a sufficient number of videos of different individuals and are inadequate for our applications. We collected our own database. Table1 describes the frame parameters.

TABLE I. PARAMETERS OF THE PROPOSED TRAINING AND TESTING IMAGE DATABASE

Parameter	Value
Number of males	5
Minimum image resolution	243×360
Total number of images	2255
Number of females	5
Color space format	RGB
People's age	8–60 years old
Lighting conditions, background, and facial	No restrictions

expression	
------------	--

This dataset was split into two independent video sets: training and testing. The training set was used for feature generation and TDNN classifier construction. A performance evaluation of the trained classifier was approved with the use of the testing set.

Various batch training algorithms can train a network. Three types of training algorithms have eight training functions: gradient descent algorithms (traingd, traingdm, trainrp), conjugate gradient algorithms (traincg, traingcg, traingcp), and quasi-Newton algorithms (trainbfg, trainlm) according to the explanation in [25]. dependent on the performance results, we used quasi-Newton algorithms (trainbfg). Determining an appropriate NN architecture depends greatly on the experience of the person controlling the experiment [16]. The feature vector for gender recognition by the network has to be encoded and placed in the input nodes of the network. These input values must be scaled so they are in the same range which can help save time and get high performance of the training methods [16].

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the gender classification algorithm, we prepared our own video sample for the subject. Figure 3 shows sample face sequences from the dataset. TDNN with conjugate gradient algorithms, BFGS quasi-Newton backpropagation (trainbfg) training function in our experiments, The figure also shows the performance result of the human identification and gender and age classification system.



Fig. 3. Face Sequences from the Dataset

All videos in the experiments have a range of 100–240

Application Type	Training Set Classification Rate	Testing Set classification Rate
Gender classification		91.2
Age classification	100%	88.0
Human identification		83.0

frames. We collected 10 video samples from each person and used MATLAB 10 to implement the development process of gender classification. For each frame, global geometric features from the face segment part were extracted to form the face vector trajectories. These trajectories include vertical center, horizontal center, semi-major axis, semi-minor axis, and four-

quadrant inverse tangent (arctangent) of the real parts of Y and X for each image sequence.

The frames that did not display a face part from the samples were eliminated. The feature vector of all frames form an input vector that trains the TDNN, which takes 70% of the samples for training and 30% for testing of all persons. During this research, the TDNN architecture was applied as shown in Figure. 4. Based on previous research [16], the sigma mode function “tansig” is used in the input and hidden layers, whereas the linear transfer function pure linear “purelin” is used in the output layer. We created 4 layers of TDNN with 5 input neurons and 20 hidden neurons, with the delay of 0–3 for the first two layers and 0–5 for another layer.

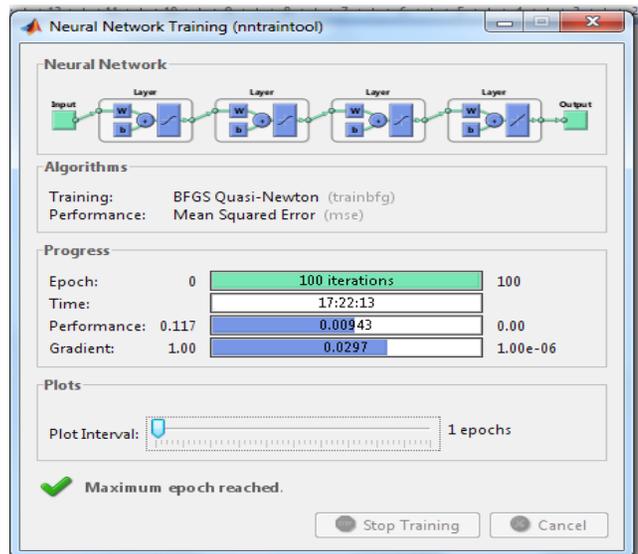


Fig. 4. TDNN Architecture

The performance results with 200 epochs of gender classification, human identification and age classification are shown in Figure.5a, b, and c, respectively.

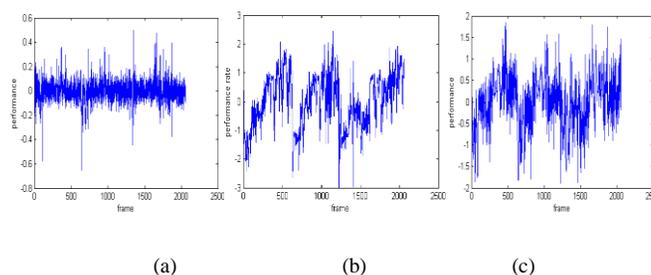


Fig. 5. Mean Squared Error of (a) Gender Classification, (b) Human Identification, and (c) Age Classification

TABLE II. CLASSIFICATION RATE USING 70% FOR TRAINING AND 30% FOR TESTING

Table 2 lists the experiments results based on classification rate and application purposes. The conjugate gradient algorithms (trainbfg) achieved promising results in training the TDNN [16]. The results were sorted based on the time consumed and the classification rate, where the highest classification rates in testing sets are in gender classification,

age classification, and human identification. This result implies that gender, age, and human beings can be classified with promising recognition rate by using global geometric features from video. The developed method based on global geometric facial features achieves a high classification rate of 100% in the training set for all application, as well as 91.2% for gender classification, 88% for age classification, and 83% for human identification in the testing set. The proposed method also establishes the real-time system to be used in three applications with a simple computation for feature extraction. During the experiment, we attempted to train 30 frames from each sample and achieved the same result from all frame trainings. The results show that the suggested methods with all frames and with a sample sequence of frames outperform state-of-the-art methods.

V. CONCLUSION AND FUTURE WORK

A real-time classification system using TDNN is described according to BFGS training algorithms and used for gender and age classification and human identification. The developed system implies that we can classify gender, age, and human being with promising recognition rate by using global geometric features from video. The developed method based on global geometric facial features achieves a high classification rate of 100% in the training set for all application, as well as 91.2% for gender classification, 88% for age classification, and 83% for human identification in the testing set. Future work will be devoted to implementing and examining the other features extracted from face parts by various techniques with real-time application.

ACKNOWLEDGMENT

This research was funded by the Deanship of Scientific Research in Zarqa University, Jordan.

REFERENCES

- [1] Alok Verma, S Achyut Rajy, Abhishek Midyaz, Jayasree Chakraborty,(2014). Face Detection using Skin Color Modeling and Geometric Feature. Conference: 2014 International Conference on Informatics, Electronics & Vision (ICI)). Dhaka: IEEE, 2, 14.Pp. 1-6. <http://dx.doi.org/10.1109/ICIEV.2014.6850755>
- [2] Erno Mäkinen, Roope Raisamo. (2008). an Experimental Comparison of Gender Classification Methods. Pattern Recognition Letters, 1544–1556. <http://dx.doi.org/10.1016/j.patrec.2008.03.016>
- [3] Md. Hafizur Rahman, Suman Chowdhury, Md. Abul Bashar. (2013). an Automatic Face Detection and Gender Classification from Color Images using Support Vector Machine. Journal of Emerging Trends in Computing and Information Sciences, 5-11.
- [4] Meltem Demirkus, Matthew Toews, James J. Clark, Tal Arbel. (2010). Gender Classification from Unconstrained Video Sequences. IEEE, 55-62. <http://dx.doi.org/10.1109/cvprw.2010.5543829>
- [5] Ms. Rutuja G. Shelke, Prof. S. A. Annadate. (2014). Face Recognition and Gender Classification using Features of Lips. International Journal of Innovation and Scientific Research, 337-341.
- [6] Muhammad Naem Ahmed Khan, Sheraz Ahmed Qureshi and Naveed Riaz. (2013). Gender Classification With Decision Tree. International Journal of Signal Processing, Image Processing and Pattern Recognition, 165-176.
- [7] Prashanth Kumar G. And Shashidhara M. (2014). Real Time Detection and Tracking of Human Face using Skin Color Segmentation and Region Properties. International Journal of Signal Processing Systems, 102-107.
- [8] Ramesha K, K B Raja, Venugopal K R And L M Patnaik. (2010). Feature Extraction Based Face Recognition, Gender and Age Classification. International Journal on Computer Science and Engineering, 14-23.
- [9] Reecha Sharma, M.S. Patterh. (2015). Indian Face Age Database: a Database for Face Recognition with Age Variation. International Journal of Computer Applications (0975 – 8887), 21-27.
- [10] Reza Hassanpour, Asadollah Shahbahrani, and Stephan Wong. (2008). Adaptive Gaussian Mixture Model for Skin Color Segmentation. Proceedings of World Academy of Science, Engineering and Technology, 1-6.
- [11] Sajid Ali Khan, Muhammad Nazir, Naveed RIAZ. (2013). Optimized Features Selection for Gender Classification using Optimization Algorithms. Turkish Journal of Electrical Engineering and Computer Sciences, 1479 - 1494. <http://dx.doi.org/10.3906/elk-1203-51>
- [12] Vladimir Khryashchev, Lev Shmaglit, Andrey Shemyakov, Anton Lebedev, (2014). Gender Classification for Real-Time Audience Analysis System. Open Innovations Association FRUCT, Proceedings Of 15th Conference of Fruct Association, St. Petersburg: IEEE, 2014, Pp. 52 - 59. <http://dx.doi.org/10.1109/fruct.2014.6872428>
- [13] Daoud, E. A. (2015). Face Recognition Fuzzy Clustering and Kernel Least Square. Journal of Computer and Communications, 2015, 1-7. <http://dx.doi.org/10.4236/jcc.2015.33001>
- [14] Ahmad Awwad, Ashraf Ahmad, Walid A. Salameh. (2013). Arabic Race Classification of Face Images. International Journal of Computers and Technology, 234-239.
- [15] Akshay Kajekar, Shreyash Patel, Sweta Chauhan, Prashasti Kanikar. (2012). Blending of Face Segments. International Journal of Computer Science and Information Technologies, 3929-3933.
- [16] AL Mashagba, E. F. (2015). Human Identification by Gait Using Time Delay Neural Networks. Computer and Information Science, 56-63. <http://dx.doi.org/10.5539/cis.v8n4p56>
- [17] Anushri Jaswante. Asif Ullah Khan. Bhupesh Gour. (2013). Gender Classification Technique Based on Facial Features using Neural Network. International Journal of Computer Science and Information Technologies, 839-843.
- [18] Chuanxu, W. (2011). Face Segmentation Based on Skin Color in Complicated Background and its Sex Recognition. Journal of Software, 1209-1216.
- [19] Dewi Agushinta R., Adang Suhendra, Sarifuddin Madenda, Suryadi H.S. (2010). Face Component Extraction using Segmentation Method on Face Recognition System. Journal of Emerging Trend in Computing and Information Sciences, 67-72.
- [20] Federico Matta, Usman Saeed, Caroline Mallauran and Jean-Luc Dugelay, (2008) Facial Gender Recognition using Multiple Sources of Visual Information. Multimedia Signal Processing, 2008 IEEE 10th Workshop On, Cairns, Qld: IEEE, 2008, Pp. 785-790. <http://dx.doi.org/10.1109/MMSP.2008.4665181>
- [21] Hlaing Htake Khaung Tin. (2012). Perceived Gender Classification From Face Images. I.J. Modern Education and Computer Science, 8-12.
- [22] Islaml, M. S. (2013). Gender Classification using Gradient Direction Pattern. Sci. Int (Lahore), 797-799.
- [23] Jawad Nagi, Syed Khaleel Ahmed, Farrukh Nagi. (2008). A MATLAB Based Face Recognition System using Image Processing and Neural Networks. 4th International Colloquium on Signal Processing and its Applications, 83-88.
- [24] Md. Hafizur Rahman, Tonmoy Das, Manamatha Sarnaker. (August 2013). Face Detection and Sex Identification from Color Images using Adaboost with SVM Based Component Classifier. International Journal of Computer Applications, 0975 – 8887.
- [25] Sanjeev Sharma, Ritu Tiwari, Anupam Shukla And Vikas Singh. (2011). Identification of People using Gait Biometrics. International Journal of Machine Learning and Computing, Vol. 1 (No. 4), 409-415.
- [26] Swaminathan, A. (2000). Gender Classification using Support Vector Machines., 2000. ENEE633: Statistical and Neural Networks pattern Recognition, 2000, (Pp. 306 - 311). <http://dx.doi.org/10.7763/IJMLC.2011.V1.61>
- [27] Xiaodong Li, Wei Jiang, Weijie Mao, Jian Li And Kai Chen. (2013). Gender Classification Based on Evolutionary Extreme. International Journal of Innovative Computing, Information and Control, 3839-3849.
- [28] Young, Cynthia Y. (2010). "Chapter 9". Precalculus. John Wiley and Sons. p. 831. ISBN 0-471-75684-9